DEVELOPMENT OF MULTIVARIATE PREVALENT CANCER MAPPING MODEL IN LIBYA

MARYAM AHMED SALEM ALRAMAH

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

UNIVERSITI PENDIDIKAN SULTAN IDRIS

2018

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

DEVELOPMENT OF MULTIVARIATE PREVALENT CANCER MAPPING MODEL IN LIBYA

MARYAM AHMED SALEM ALRAMAH

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (APPLIED STATISTICS)

FACULTY OF SCIENCE AND MATHEMATICS
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2018

05-4506832 pustaka.upsi.edu.my Perpustakaan Tuanku Bainun Kampus Sultan Abdul Jalil Shah PustakaTBainun ptbupsi

Please tick (√):

Project Paper	
Masters by Research	
Masters by Mix Mode	
Ph.D.	√

INSTITUTE OF GRADUATE STUDIES

DECLARATION OF ORIGINAL WORK

This declaration is made on the 17th July 2018.

i. Student's Declaration:

I, Maryam Ahmed Salem Alramah, P20122002014, Faculty of Science and Mathematics hereby declare that the work entitled (DEVELOPMENT OF MULTIVARIATE PREVALENT CANCER MAPPING MODEL IN LIBYA) is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

_____

Signature of the student

ii. Supervisor's Declaration:

I Nor Azah binti Samat hereby certifies that the work entitled (DEVELOPMENT OF MULTIVARIATE PREVALENT CANCER MAPPING MODEL IN LIBYA) was prepared by the above named student, and was submitted to the Institute of Graduate Studies as a * partial/full fulfillment for the conferment of DOCTOR OF PHILOSOPHY, and the aforementioned work, to the best of my knowledge, is the said student's work.

_____ _____

Date Signature of the Supervisor

ACKNOWLEDGMENTS

First and foremost, all praised to the Almighty Allah (God) who has been very kind for giving me this chance to pursue my study programme.

This Ph.D. would not have been possible without significant contributions and assistance from a number of people. Above all else, I would like to thank my research supervisor Associate Professor. Dr. Nor Azah Samat, who came up with the proposal for me to do this Ph.D. I am indebted to her for many sensible suggestions and ideas and above all for many discussions which benefited from her uniquely simultaneous insights into Bayesian disease mapping, statistical modelling, and spatial epidemiology. Also, she suggested methods for the analysis, verified the results, and proofread numerous drafts. So, I might want to express gratitude toward her for all guidance; for the time, she has given to this study; for the help that gave personal time to work; and the greater part of all, to make me certain with my work. Without her, this thesis would not have been conceivable.

Similarly, I am very indebted to my co-supervisor, Associate Professor. Dr. Zulkifley Mohamed, the Dean of Faculty of Science and Mathematics, UPSI, who showed me what I needed to think about Applied statistic, and who managed to come up with another idea if a specific technique did not appear to work. His constancy and his knowledge of spatial methods have majorly affected a few studies in this thesis and I am extremely thankful for this. I would like to show my appreciation to examining committee members, for carefully reading the dissertation and for their suggestions on improving it.

Most importantly, I would like to thank my parents, Ahmed Salem and Amna Mohamed, both of you are precious gifts from God, and my siblings, my sisters, Fatmah and Hanan, my brothers Esam and Salem for their love and support and I am especially thankful to my kids, Lujain, Mohamed, Gumana and Jood, for being loving and supportive, believing in me and always being there for me. I'm proud of everything we have achieved together. I owe my loving thanks to my lovely husband, Firas Mohamed. I thank God for having him as my husband, best friend, and soul companion. Without his encouragement and understanding it would have been impossible for me to finish this work.

I deeply thank for the Ministry of Higher Education, Libya, for funding the Graduate Scholarship, is gratefully acknowledged. Also, I would like to thank the university of Tripoli for giving me the opportunity to study towards my PhD and all the amazing people that work there. In particular, my father-in-law, Professor. Dr. Mohamed Shallouf, for being a great friend and for helping me to get cancer data for this research from African Oncology Institute, that enabled me to do the analysis and subsequently write this PhD. I would also like to thank the Department of Statistics, Faculty of Science at University of Tripoli, Libya for their training and encouragement in the field of statistics.

Lastly, five years' Ph.D. study in the Department of Mathematics at Universiti Pendidikan Sultan Idris, Malaysia is great experience to me. So, I warmly thank all members and staffs that work in.

ABSTRACT

The aim of this study was to develop the best possible method for providing estimates of relative risk for cancers by utilizing a multivariate approach through shared component model to produce specific and shared spatial patterns of cancer disease in Libya. The methodology in this study is statistical model to development to jointly analyze the spatial variation in rates of several cancers with their four spatial components without any conditions. The data for incidence of breast, lung, prostate, bladder, stomach, colon and liver in Libya districts, were used in this study. Four risk factors representing smoking, low physical activity, overweight and alcohol consumption were considered. A general framework for jointly modelling the variation of seven cancers, some of which share latent spatial fields, is considered. Finally, a new Multivariate Shared Component model is successfully developed. The procedure for predicting relative risks based on the model is presented in order to suggest approximate future cancers diseases progression patterns for Libya. In this study, the new model and the existing models were then applied to the morbidity data of seven types of cancer and four components for female and males in Libya during 2006-2011. An overall assessment of the models was carried out based on the goodness of fit. The results of the Deviance Information Criterion (DIC) value (BYM: 821.841, Mixture model: 797.832, SCM: 794.788) demonstrated how the new model fits the data better than the existing models. Furthermore, the findings of the analysis or the spatial maps analysis showed that on the basis of the model, the areas of high risk were relatively distributed across the region. As a conclusion, an observed comparison of the existing models and the new model based on the posterior precision of the relative risk estimates, strongly suggests that the multivariate modelling approach is a valuable expansion over individual analyses. The shared component model also can be readily applied to other; cancers or non-communicable diseases, risk factors, spatial areas, and time periods. The implications is that the model will provide accurate and clear maps of cancer morbidity across all regions in Libya. The information will be valuable for guiding public health strategies to reduce cancer risk, as well as of risk -lifestyle behaviours through cancer disease and maintaining a healthy lifestyle.

PEMBANGUNAN MODEL MULTIVARIAT BAGI PEMETAAN KANSER DI LIBYA

ABSTRAK

Kajian ini bertujuan membangunkan kaedah terbaik bagi menganggar risiko relatif kanser dengan menggunakan pendekatan multivariat melalui model komponen terkongsi untuk menghasilkan corak ruangan khusus dan terkongsi penyakit kanser di Libya. Metodologi dalam kajian ini ialah perbincangan model statistik bagi menganalisis variasi ruangan secara bersama terhadap kadar penyakit kanser berserta empat komponen ruangan yang berkaitan tanpa sebarang syarat. Data berkaitan insiden kanser payudara, prostat, pundi kencing, perut, usus besar, dan hati yang berlaku di daerah-daerah di Libya digunakan dalam kajian ini. Empat faktor risiko diambil kira yang terdiri daripada merokok, kekurangan aktiviti fizikal, berat badan berlebihan, dan pengambilan alkohol. Rangka kerja umum bagi pemodelan bersama variasi tujuh penyakit kanser diambil kira dan sebahagian daripadanya berkongsi kawasan ruangan terpendam. Akhirnya, model baharu komponen Terkongsi Multivariat telah dibangunkan. Prosedur untuk meramal risiko relatif berdasarkan model terkini diperkenalkan bagi memberikan anggaran corak perkembangan penyakit kanser di Libya. Dalam kajian ini, model baharu dan model sedia ada dipadankan terhadap data morbiditi tujuh penyakit kanser dan empat faktor risiko untuk wanita dan lelaki di Libya pada 2006-2011. Penilaian keseluruhan terhadap model dilakukan berdasarkan ketepatan padanan. Hasil dapatan nilai *Deviance Information Criterion* (DIC) (BYM: 821.842, *Mixture model*: 797.832, SCM: 794.788) menunjukkan model baharu yang diperkenalkan mempunyai ketepatan padanan data yang lebih baik daripada model-model umum yang dikaji. Seterusnya, hasil dapatan analisis peta ruangan berdasarkan model yang sedia ada menunjukkan kawasan-kawasan yang berisiko tinggi tertabur secara relatif di kawasan tersebut. Sebagai kesimpulan, perbandingan empirik model-baharu dan model sedia ada berdasarkan krjituan posterior bagi anggaran risiko relatif mencadangkan bahawa pendekatan pemodelan multivariat adalah lebih berguna dan bernilai berbanding analisis secara individu. Model komponen terkongsi multivariat juga boleh digunakan bagi pelbagai jenis kanser atau penyakit tidak berjangkit yang lain, faktor risiko, kawasan ruangan, dan tempoh masa. Implikasinya, model ini menyediakan pemetaan morbiditi penyakit kanser di seluruh Libya dengan lebih tepat dan jelas. Maklumat tersebut adalah sangat bernilai untuk membantu melaksanakan strategi kesihatan umum bagi mengurangkan risiko penyakit kanser serta risiko sikap gaya hidup melalui penyakit kanser, dan mengekalkan gaya hidup yang sihat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AIC	Akaike Information Criterion
AOI	African Oncology Institute
BMI	Body Mass Index
BYM	Besag, York And Mollie Model
CAR	Conditional Autoregressive
CD	Cancer Disease
DIC	Deviance Information Criterion
GIS	Geographic Information System
JDM	Joint Disease Mapping
MCMC	Markov Chain Monte Carlo
MHL	Ministry of Health Libya
MVDM	Multivariate Disease Mapping
MVSCM	Multivariate Shared Component Model
NCDs	Non- Communicable Diseases
RR	Relative Risk
SCM	Shared Component Model
SIR	Standardized Incidence Ratio
SMR	Standardized Mortality or Morbidity Ratio

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

The mapping of disease incidence and its prevalence has provided and achieved huge benefits in the public health sector, in parts of, epidemiological studies and the studies of human diseases (Koch, 2005). The rapid evolution of life has led to the emergence of several factors that affect the spread and outbreaks of the disease case, these factors are like booming population, rapid urbanization, environmental pollution and global warming. Studies of the disease have confirmed that knowledge of the disease case location and its outbreak is one of the important reasons to help the government's policy and financial support to understand the strategy of the disease.

Currently, there are some factors, which consider as causes of disease, as mentioned earlier. All of these latest disease factors have influenced the conditions for disease outbreaks in general and cancer disease in particular. Exploring diseases affected areas can reveal many spatial aspects of the disease. This is known as "disease mapping", which helps the public to understand causes of disease and some puzzles of disease outbreak. Because it presents a optical means of identifying cause and effect relationships existing between humans and their environment. In addition, by using spatial mapping aspects of diseases, it can help people understand some puzzles of a disease outbreak and it can be seen from the map of disease more than a disease by using joint disease mapping or multivariate disease mapping.

The representation and analysis of geographical distributions of disease are widely studied by biostatisticians, epidemiologists, medical demographers and academics. Moreover, it has become an important technique in disease epidemiology and health research. According to Barrett (2000), a world map of diseases was produced in 1792, by a German physician named Finke. This is much earlier than any world disease map previously known. However, since it was a manuscript map that could not be re-produced, Barrett (1993) concluded that the original has been either lost or destroyed. This is based on the results of investigations done by him that there is no subsequent reference to Finke's map (Barrett, 1993). On the other hand, according to Barrett (2000), contemporary researchers believe that either the Berghaus map produced in 1848 or the Schnurrer map produced in 1827, is the first world disease map.

The term "disease mapping" is a method of research that is commonly used to highlight the geographical variability of disease on maps by using different colours to differentiate the risk levels in the certain area.

Disease mapping is widely used in descriptive epidemiology to display morbidity or mortality information of an area. When, the term " morbidity" refers to incidence and "mortality" refers to death. Statistical and epidemiological studies have proven that the disease mapping method is better than the raw disease (Berke, 2004). The use of disease mapping is very important to identify areas with a high risk that deserve closer scrutiny and attention compared with other regions. However, the production of disease maps relies on a modelling to estimate and predict the risks. Therefore, better prediction and careful estimation of the risk, subsequently would produce accurate maps (Knorr-held, Raler, & Becker, 2002).

The core objective of disease mapping is the relative risk estimation. Knorr-Held et al. (2002) has found that the estimation of cancer disease risk is considered as the first order in each country, especially in the developing countries. Most studies that estimated relative risk, involve exploratory analysis. These studies focused on covariates and their effects on disease geographical distribution and the study of geographic information systems (GIS) to integrate patient-related information. Nevertheless, no previous study has examined the model proposed by Held, Natário, Fenton, Rue and Becker (2005). Therefore, to further investigate cancer disease mapping, the present study attempts to introduce an alternative method of relative risk estimation based on the joint models or multivariate

models for the cancer diseases, initially based on the model proposed by Held et al. (2005). The proposed alternative model will help to develop procedures for predicting the disease risk for future times based on the results of cancer map and the observed history.

The primary aim of this research is to study the geographical distribution of cancer disease in Libya. According to Lawson (2001), there are three main areas of application in disease study, which have used in the field of representational disease study. These are as follows:

1) Disease clustering: the purpose of this analysis is to examine potential environmental hazards based on localized clusters of cases. It is also used to determine the locations of any such clusters, which are important in public health surveillance.
2) Disease mapping: the purpose of this method is to estimate the true relative risk of a disease across a geographical area. The primary objective is to reduce the noise in a disease map.
3) Ecological analysis: it focuses on the analysis of geographical distributions of disease in relation to explanatory variables, where several supplementary issues relating to disease mapping can be specified.

In the case of cancer disease, there were not many studies that used multiple statistical methods to estimate the relative risk for disease mapping (Knorr-Held et al., 2002; Held et al., 2005; Mahaki et al., 2011). Most cancer studies were based only on a

single analysis of a particular cancer disease, which includes the study of spatial variations in disease rate (see, for example, Torabi & Rosychuk, 2011; López-Abente, Aragonés, García-Pérez, & Fernández-Navarro, 2014). However, there are only few studies that used joint disease models for disease mapping (see, for example, Downing, Forman, Gilthorpe, Edwards, & Manda, 2008). Some of them used the geographic information system (GIS) to incorporate the patient-related information of dengue disease (Mohd Din, Shaaban, Norlaila, & Norariza, 2007) and of cancer disease (Elebead, Hamid, Hilmi, and Galal, 2012). A geographic information system (GIS) is a computer database management system that is used to provide an effective method of organizing, storing, mapping and displaying the large amounts of information in the question.

Cancer is recognized as a growing health problem, particularly in the Eastern Mediterranean Region (EMR), that includes Libya as well. Cancer is considered one of the most common health problems in Libyan populations, both in female and male (Abussa, 2007; Mussa et al., 2015). Due to some social reasons, most of the cases come very late for a medical examination. Therefore, most of the medical reports have shown this problem is increasing every year. At the same time, this brings the attention of health authorities. Most of the cancer patients are diagnosed in the last stages of the disease. Hence, there is an urgent need to evaluate the factors, which are influencing the increase of cancer in Libya.

One of the effective approachs to preventing this disease is to understand the epidemiology and ethology of the problem. The present study deals with cancer disease in Libya to determine the areas with high and low risk compared with other areas. There are

so many cancer cases in Libya and it keeps on increasing every year, according to the annual reports from cancer registry centres in Libya and previous studies (Abussa, 2006, 2007; Mussa et al., 2015).

The current approach of cancer mapping in Libya is based on the primitive method of counting total numbers of disease occurrences in every region. For example, higher occurrence of disease is estimated when conventional approaches are employed to determine huge amount of cases in particular areas. This method does not take into account some important factors, such as the population size and the land size of individual regions. From a global perspective, prevention and control strategies for cancer disease comprise of disease surveillance and control, education, inter-agency collaboration, community participation and quality assurance. Nevertheless, these methods are effective only after the cancer cases have occurred. In Libya, there is no proper system of national cancer surveillance and integrated care for non-communicable disease prevention. In other hand, there is mostly lack of national cancer surveillance, also a model of integrated care for non-communicable disease prevention, cancer in general, and inadequate national capacity building and program sustainability.

Therefore, in order to prevent these diseases before occurrence, the current research aims to introduce an alternative method of predicting the high-low risk areas based on the multivariate models for cancer disease, which takes into account significant covariates. In other words, in this study, it is interesting to know and outline the current burden of cancer in the region and the challenges which are faced by the Libyan Government.

The African Oncology Institute (AOI), published recently annual reports of morbidity in Libya (Abusaa, 2011). Population and morbidity information were collected for the periods 2006/2008 to 2011/2012, to identify differences in morbidity of disease in Libya. In precise, this study had gathered data from 2006 until 2011, a total of 6 years, in which data had been retrieved from the AOI Department of Statistics. This study will highlight regions with high risk that need urgent actions. The study also suggested for further research to introduce a better spatial model for estimating cancer disease in Libya.

Regarding of cancer in Libya, recent researches have shown that breast cancer is one of the most prevalent cancers among Libyan women (See, for example, BenNasir, El Mistiri, McGowan, & Katz, 2015; Boder, Abdalla, Elfagieh, & Collan, 2013; Boder, Abdalla, Elfagieh, Buhmeida, & Collan, 2013; Boder et al., 2011; Bodalal, Azzuz, & Bendardaf, 2014; Bodalal & Bendardaf, 2014; Ermiah, 2013). On account of the burden cancer in Libya, the main source of information for this study that collected by AOI was the initial motivation to attempt for finding better statistical models or spatial modelling methods for estimating cancer disease in Libya.

With this background, the current study will introduce multivariate models for cancer disease, considering covariates and risk factors, such as smoking, low physical activity, overweight and alcohol consumption. Application of these models will provide accurate and clear maps of cancer morbidity across all regions in Libya. The information will be valuable for guiding public health strategies to reduce cancer risk, as well as of risk-lifestyle behaviours through cancer disease and maintaining a healthy lifestyle.

This work introduces the first comprehensive study of some types of cancer in Libya, using all Libyan districts. Based on the collected data from the African Oncology Institute (AOI), Sabratha, this research adds a geographical significant segment to the recent ongoing evaluation of the cancer incidence in Libya.

Then, the problem statement of the current research is explained and discussed in next section.

1.2 Problem Statement

Libya is one of the countries in Africa (Arab countries) with a high rate of cancer disease. According to WHO report on cancer (2014), cancer diseases increased significantly and it is ranked as the major cause of death in Libya (Stewart & Wild, 2014). Cancer threatens human welfare, socio-economic advances, productivity and social cohesion. Cancer prevention by understanding the proper ways to maintain a healthy lifestyle contributes directly to the prevention of this deadly disease.

Recently, the statistical studies have provided remarkable results in all fields and played an important role in policy formulation. Therefore, the statistical applications or applications of statistics are ubiquitous in all areas of life, such as in business, civil engineering, real life, setting an insurance rate, decision making, etc. In statistics field, statistics is defined as a discipline that is concerned with the treatment of numerical data,

which is derived from groups of individuals. These individuals often are members or people, such as those living together in a certain area or those infected with a particular disease, also they may include animals or other organisms.

In recent years, there has been considerable interest in the development of spatial methods to analysis spatially correlated data, which are often geographically referenced, temporally correlated or joint (multivariate) for disease or a number of diseases. The literature shows that there is a lack of risk assessment of cancer in Libya. Moreover, the publications that consider the geographical variability or the estimation of relative risk in the use of spatial models for cancer disease in Libya at all, yet, this epidemiological phenomenon has not been studied.

In Libya, the current methods for high and low-risk areas are still based on the total numbers of disease occurrences across the regions, without considering other factors, such as the population size of individual areas. Therefore, this research is interested in introducing an alternative method in order to estimate and predict the high-low risk areas, based on the multiple models for a set of cancer disease.

In the literature, there has been a rapid growth in disease mapping methods over the past twenty years. Despite the rapid progress in this area, there are still some key problems that are not addressed adequately in the current literature. The present study will attempt to identify such problems and recommend solutions.

Disease mapping refers to a set of statistical techniques that leads to provision of maps based on estimation of disease rates. In the field of disease mapping, joint analysis of diseases has spelt several potential benefits, such as enhancing the epidemiological interpretability and improving the precision of estimates, over mapping a single disease. In addition, spatial analysis for multiple diseases suggests additional benefits over spatial analysis meant for a single disease. Hence, due to its vast advantages, the multivariate disease mapping technique has been developed and investigated substantially in recent years, whereby several methods have been proposed to map two or more diseases jointly. These methods employ a particular risk factor or shared component that emerges as the primary for all common cancers. In addition, there is direct knowledge of them.

Among them, the shared component model has appeared popular in these recent years for its wide application in a substantial number of studies. Nevertheless, this model has yet to be incorporated with joint analysis that probes into several diseases, along with their risk factors, with absence of condition, in both estimating and predicting the relative risk. As such, this thesis purports the extension of this model by embedding the multivariate spatial among diseases for relative risk and to predict relative risk to suggest approximate future cancer disease progression patterns in Libya. Due to the rising morbidity and infirmities recorded in Libya due to cancer, this study has attempted to model the incidence rates of seven prevalent cancers and four related components to represent their related risk factors, particularly in Libya, by employing the spatial shared component approach.

Therefore, this thesis aims to develop a new approach based on the model of the multivariate shared component model or shared component model and the use of joint approach for a set of cancer disease to estimate the relative risk (the relative cancer risk). It attempts to model the geographical variation of the certain disease, especially cancer disease using a hierarchical Bayesian approach, in order to investigate the different sources, which affected the estimation of relative risk and its methods. This study presents four various types of hierarchical models related to mapping of disease, hence allowing the inclusion of Bayesian method to determine the effects sought. This study also reported the results of fitting these models to the cancer data in Libya, after taking into consideration the spatial autocorrelation.

Finally, the shared component model was fitted so as to identify the variation in the relative risk within the different regions of the country. To the best of our knowledge, analyses of multiple geographical patterns for several cancers diseases have not been done in Libya. Wordwide, except a study conducted by Mahaki and colleagues that used seven cancers in Iran but the human development index (HDI) as a socioeconomic factor has been consided utilizing as common risk factors for all the selected cancers with their direct knowledge (Mahaki et al., 2011). In other words, no published work using proposed SCMs to address analyzing multiple diseases jointly related to Libyan geographical and health sectors. The purpose of this study is (a) to quantify the essential risk factors of cancer and (b) to produce spatial maps using a model of Bayesian spatial, jointly.

The following section contains a brief explanation of the aims or objectives of the current study.

1.3 Research Objectives

Given the high prevalence of cancer in Libya, it is very important to investigate the spatial patterns of cancer disease as preventative measures. This study aims to highlight spatial disparities in coverage of a set of cancers in Libya. Throughout this research, we will address the following research objectives to improve the relative risk estimates of multiple cancers disease instead of a single cancer disease morbidity or incidence. For this purpose, the current research will address the following research objectives:

1) To determine the spatial patterns and risk factors of seven types of cancer in Libya for the period of six years from 2006 to 2011,
2) To propose an improved shared component model (SCM) of relative risk estimation for mapping cancer disease,
3) To compare the findings of relative risk estimation based on the shared component model with existing statistical models for disease mapping using cancer data in Libya,
4) To examine the performance of different models in estimating the relative cancer risk via monitoring the history graphs and producing maps of cancer. In addition, to produce the first base maps for implementation of Cancer Control Program in Libya for seven types of cancers, individual and joint, respectively,

5) To develop procedures for predicting the disease risk based on shared component model (SCM).

1.4 Significance of the Research

This research is concerned about seven types of cancer diseases in Libya which are breast, lung, prostate, bladder, stomach, colon and liver. As a matter of fact, this study has associated these cases of cancer with four prominent risk factors, which are: 1) sedentary lifestyle, 2) being overweight, 3) the habit of smoking, and 4) excessive consumption of alcohol. In other words, because there are clear relationships between each other, which called risk factors or lifestyle factors, this study is considered about a set of cancers with related shared components. As such, in the end of this thesis, the reader may differentiate the types and symptoms of selected cancers diseases. The significance of this study in Libya is to identif similar patterns for multiple cancer diseases, because it will provide stronger evidence than that of separate analyses for single cancer. The information will be useful for public health strategies to reduce cancer risk, because of its important in public health surveillance, the ease of interpretation, as well as ability to identify shared and specific patterns of risk among different diseases. For people, by utilize joint analyses of spatial data, it might be possible to better understand the relationship between personal disease risk and the geographic distribution of disease risk.

The previous epidemiologic studies (see, for example, Dayhum et al., 2017; De Araújo et al., 2013; Diah, Aziz, & Ahmad, 2016) had discussed mainly the spatial analysis of a single disease to estimate the relative risk, without considering other factors. Only a few publications considered the estimation of relative risk for a set of disease that shares common risk factors, that havd been considered a main shared component for all selected diseases (see for example, Held et al., 2005; Downing et al., 2008; Mahaki et al., 2011). Therefore, this research focuses on seven cancers diseases in Libya, theier names as mentioned early breast, lung, prostate, bladder, stomach, colon and liver cancer, where the characteristics of all diseases are discussed in Chapter 2. Among these seven cancers, there are some similarities and differences, which will be outlined at the end of the research.

This study will produce disease maps that will highlight the different levels of risk, such as the very high, high, low and very low-risk regions of disease occurrences. Consequently, these maps will support the responsible authorities in the country to identify risk levels, which require further scrutiny and intense attention. Of the benefits of production of these maps, the governments may able to prepare the medical requirements and the awareness necessary to identify the sources of hazards related to the disease. In addition, for the citizen, the use of maps of disease by specific communities provides them with knowledge of whether there are areas with high risks or low risks of disease occurrences.

One of the significant aspects of such maps is to highlight the spread of more than one disease at the same time in one area. Since, the relationship between the certain type

of cancer and their risk factors is a fact, these maps can display the geographical distributions of the disease with their factors. Because, multivariate disease mapping of cancer disease refers to the joint mapping of multiple cancers diseases from regionally aggregated data and continues to be the subject of considerable attention for biostatisticians and spatial epidemiologists, as well as for individuals. The key issue in multivariate cancer disease is to map multiple cancers disease correlations among themselves.

Joint or multivariate disease mapping, which may be defined as the joint spatial analysis of several diseases, is one of the most widespread approaches to estimation of the relative risk of diseases jointly, with the assumption that the multiple diseases are independent. The main goals of the shared component model (SCM) are to: interpret, develop in the precision of the underlying diseases pattern estimation, describe the geographic variation of diseases risk, able to identify shared and specific patterns of risk among different diseases, and improvement in model evaluation criteria. Models with significant parameter estimates may still be considered invalid in a predictive sense. In this research, predictive relative risks and diseases maps based on the SCM are proposed and suggested. The posterior expected relative risks are believed to produce more accurate estimates in future time based on the proposed model. With that, the estimated disease risk maps may be helpful in identifying the high-low areas of risk for the prevalence of cancer in Libya.

In the case of Libya, these maps give us an idea of the geographical distribution of the cancer as well as cancer maps or predicted cancers maps with thier risk factors could

give a clear picture, to determine cancers density areas as a guiding component for factors causing cancers on these Libyan districts and could be used as a tool for identifying regions, which need scrutiny or further attention in term of government policy and financial support. In other words, disease maps may be useful especially for government agencies to allocate resources or identify hazards related to disease.

At the end of this research, a new improved multivariate model of cancer disease is proposed. It is expected that this improved method of relative risk estimation could produce accurate maps for cancer diseases. Finally, the spatial model that has been developed with seven cancers and their four-selected risk is new, and have never been used to model the geographical variation in incidence and risk factors of selected seven cancers in Libya. Within this context, it seems that there is no published work using the shared component model for estimating the relative risk of these cancers diseases related to the Libyan health and geographical sector. To conclude, the main objective of this research is to develop a statistical model, shared component model, based on data of multiple cancer in Libya.

The following section comprises an outline of the thesis, which gives an overview of the current study.

1.5 Organization and Scope of the Thesis

This thesis is composed of seven themed chapters as shown in Figure 1.1. The first chapter discusses the introduction of the research. It begins with a brief overview of background and motivation of the research. Problem statement and research objectives are discussed in Sections 1.2 and 1.3 respectively. Section 1.4 contains a brief description of the significance of the research.

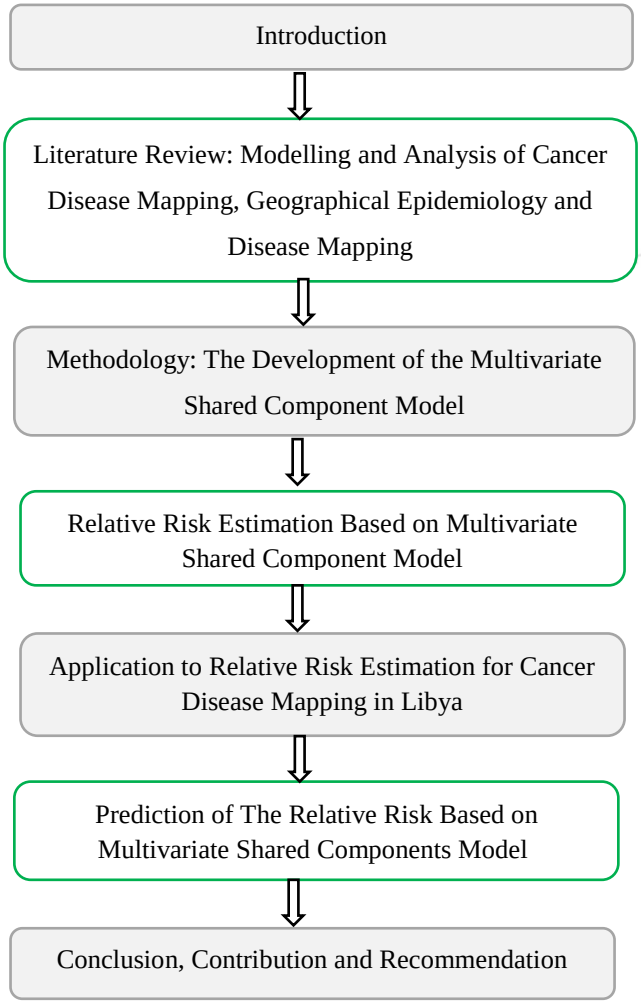

Figure 1.1. The Structure of the Thesis

Second chapter aims to review the relevant literature about modelling and analysis of cancer disease mapping. As well as the geographical epidemiology and disease mapping (DM). Section 2.1 gave introduction to organize the chapter. Then, it begins with the modelling and analysis of cancer disease mapping, in Section 2.2. This section starts by introducing the cancer disease and its burdens in the world and in Libya. Also, the researcher discusses the characteristics about these seven selected cancers with their risk factors cancers. Section 2.3 presents an extensive review of the literature related to the study. It describes the Bayesian analysis, geographical epidemiology, disease mapping and the spatial analysis; with a part of the methodology of the research. In this section, describe the background of spatial theory is discussed.

Chapter 3 focuses on model development. This chapter elaborates on the methods used in the current research. This chapter starts with an introduction in section 3.1, while associations between risk factors and cancer incidence are discussed in section 3.2. The researcher discussed some of the associations between risk factors and cancer incidence for each selected risk factor, which are considered in this study. The modelling and analysis of multivariate for disease mapping are shown in section 3.3, which including modelling disease multivariate. In Section 3.4, the researcher explores a multivariate spatial mapping for cancers. A review of the method that has been used in modelling of relative risks for two diseases and it is fit for cancer diseases. And at the end of this chapter, a section summarizes all foregoing sections, in Section 3.5.

Chapter 4 introduces an alternative method for estimating the relative risk in disease mapping, based on shared component model for several cancers, that is developed and have been discussed in Chapter 3.

Chapter 5 provides an overview of the research and its findings. In this chapter the researcher starts with an introduction. Section 5.2 focuses on data collection, description of the Libyan population and administrative boundaries. In Section 5.3, the researcher explains the application of model that is focused on relative risk estimation for cancer disease mapping in Libya. It is based the classical, namely the classical method, and on four common models, the earliest example of Bayesian mapping method. Section 5.4 describes results of the proposed model, to estimate relative risk for disease mapping to observed cancer data from all regions in Libya of six years and individual year. In this section, the results of our applications have been extracted from WinBUGS software and all maps were produced by using the GIS. Lastly, Section 5.5 given DIC value for model selection. Finally, a summary of the chapter is presented in Section 5.6.

Chapter 6 elaborates prediction of the relative risk in cancers disease pattern in Libya. A posterior predictive method is used to investigate the performance of the accuracy of the shared component model.

Finally, Chapter 7 summarizes the overall discussion and conclusions. It provides some suggestions and recommendation for further work.