THE DEVELOPMENT OF STOCHASTIC SIR-SI AGE-STRUCTURED MODEL FOR LEPTOSPIROSIS MAPPING IN MALAYSIA

AZNIDA BINTI CHE AWANG

THESIS SUBMITTED IN FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE
(MASTER BY RESEARCH)

FACULTY OF SCIENCE AND MATHEMATICS
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2018

ABSTRACT

This study aimed to develop the stochastic SIR-SI Age-Structured model to estimate the relative risk for leptospirosis mapping specifically for children and adults in Malaysia. This study used model development as it research design. The methodology of this study took into account the transmission of leptospirosis in the stochastic SIR-SI model (S=susceptible, I=infected, R=recovered for human populations and S=susceptible, I=infected for vector populations). In this study, the existing SIR-SI model was improvised and adapted to the leptospirosis transmission. Then, the model was expended to form an alternative SIR-SI Age-Structured model specifically for children and adults to estimate the relative risk of leptospirosis for these populations in Malaysia. The data used in this study were weekly data from epidemiology week 1 to epidemiology week 52 for the year 2015 for all sixteen states in Malaysia. The results of the analysis based on the age structured model were also compared with the existing models to identify the better model for estimating relative risk. For the children group, the results showed that children in Kelantan have the highest risk of contracting leptospirosis while the children in Labuan have the lowest risk of contracting the disease. Similarly, adults in Kelantan and Labuan also possessed the highest and lowest risk of contracting leptospirosis, respectively. As a conclusion, the new model was better as compared to other existing models in estimating relative risk for leptospirosis because it considered important elements such as the number of population, age group and the transmission process of the disease. This model also generates leptospirosis risk map for children and adults in Malaysia. In implication, the proposed model and generated risk maps can be practically applied towards the control of leptospirosis by government agencies, medical officers and authorities as well as increasing the awareness of local communities towards the high-low risk areas of leptospirosis.

PEMBANGUNAN MODEL STOKASTIK SIR-SI BERSTRUKTUR UMUR BAGI PEMETAAN PENYAKIT KENCING TIKUS DI MALAYSIA

ABSTRAK

Kajian ini bertujuan untuk membangunkan model stokastik SIR-SI berstruktur umur untuk menganggar risiko relatif bagi pemetaan penyakit kencing tikus terutama untuk kanak-kanak dan orang dewasa di Malaysia. Kajian ini menggunakan pembangunan model sebagai reka bentuk kajian. Metodologi kajian ini mengambil kira transmisi penyakit kencing tikus dalam model stokastik SIR-SI (S=terdedah, I=dijangkiti, R=pulih bagi populasi manusia dan S=terdedah, I=dijangkiti bagi populasi vektor). Dalam kajian ini, model SIR-SI sedia ada telah ditambah baik dan disesuaikan dengan transmisi penyakit kencing tikus. Kemudian, model tersebut telah diperluaskan bagi membentuk model alternatif SIR-SI berstruktur umur terutamanya untuk kanak-kanak dan orang dewasa untuk menganggar risiko relatif penyakit kencing tikus bagi populasi ini di Malaysia. Data yang digunakan dalam kajian ini adalah data mingguan dari minggu epidemiologi 1 hingga minggu epidemiologi 52 bagi tahun 2015 untuk kesemua enam belas negeri-negeri di Malaysia. Keputusan analisis berdasarkan model berstruktur umur juga telah dibandingkan dengan model sedia ada bagi mengenal pasti model yang lebih baik untuk menganggar risiko relatif. Bagi kumpulan kanak-kanak, keputusan menunjukkan kanak-kanak di Kelantan mempunyai risiko paling tinggi dijangkiti penyakit kencing tikus manakala kanak-kanak di Labuan mempunyai risiko paling rendah untuk dijangkiti penyakit ini. Persamaan juga dapat dilihat dalam kumpulan dewasa di Kelantan dan Labuan yang masing-masing mempunyai risiko yang paling tinggi dan paling rendah untuk dijangkiti penyakit kencing tikus. Kesimpulannya, model baru ini adalah model yang lebih baik berbanding model sedia ada yang lain dalam menganggar risiko relatif bagi penyakit kencing tikus kerana ia mengambil kira elemen penting seperti bilangan populasi, kumpulan umur dan proses transmisi penyakit. Model ini juga menghasilkan peta risiko penyakit kencing tikus bagi kanak-kanak dan orang dewasa di Malaysia. Implikasinya, model yang dicadangkan dan peta risiko yang dihasilkan boleh digunakan secara praktikal terhadap kawalan penyakit kencing tikus oleh agensi kerajaan, pegawai kesihatan dan juga pihak bertanggungjawab serta meningkatkan kesedaran masyarakat setempat terhadap kawasan berisiko tinggi dan rendah bagi penyakit kencing tikus.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DIC	Deviance Information Criterion
RR	Relative Risk
SMR	Standard Morbidity Ratio
SIR	Susceptible-Infected-Recovered for human populations
SIR-SI	Susceptible–Infected-Recovered for human populations; Susceptible-Infected for vector populations
GIS	Geographic Information System
WHO	World Health Organization

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Research Background

Disease is a pathological condition that affects the body of any living system such as humans, animals and plants particularly with the appearance of some symptoms and signs. For humans, disease can cause pain, distress, dysfunction, social problem or even death to the person affected. There are many types of diseases which are harmful to humans and one of them is infectious disease. Infectious disease is caused by pathogenic microorganisms for example bacteria, viruses or fungi and may be transmitted from one individual to another either directly or indirectly. Meanwhile, another type of infectious diseases which are harmful to humans, generally termed as zoonotic diseases, are caused and transmitted by animals. One of the most lethal diseases in this category is leptospirosis which is mainly caused by rodents and hence deserves urgency for research and development.

Leptospirosis is a worldwide zoonotic disease that affects many parts of the world including developing countries such as Malaysia. It is a disease caused by pathogenic species from the genus of Leptospira (Benacer et al., 2016) which is also known as Weil's Syndrome, Canicola Fever, Swamp Fever and so on. It is recognized as one of the world's most common re-emerging zoonotic diseases that affected many species of wild and domestic animals as well as humans. Leptospirosis is not a new disease in Malaysia because the first case of leptospirosis infection in humans was discovered by Flether in 1925 (Lim et al., 2011).

Many mammalian species have been identified as natural hosts for this disease but rodents such as rats are the main source of leptospiral infection to humans and other animals (Benacer et al., 2016). This is because the majority of human leptospirosis cases are caused by rodents (Miller, Wilson & Beran, 1991). World Health Organization (2010) recently estimated that there are more than 500 000 cases of leptospirosis each year worldwide (refer to Figure 1.1). The report also stated that most of the reported cases caused severe manifestations, for which mortality is greater than 10%.

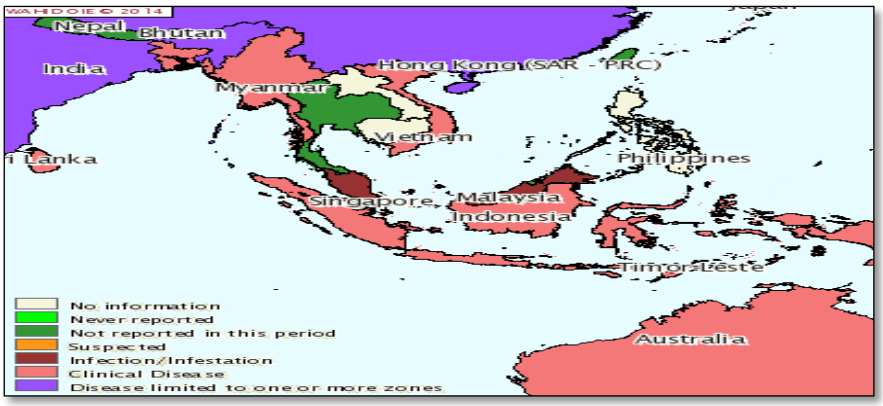

Figure 1.1. Leptospirosis distribution map in the world (WAHID-OIE, 2014)

1.2 Leptospirosis Scenario in Malaysia

In Malaysia, the Ministry of Health Malaysia (2015) in Benacer et al. (2016) highlighted leptospirosis as a mandatory notifiable disease in 2010 because the number of leptospirosis cases has risen dramatically resulting in high number of deaths. Furthermore, the number of leptospirosis cases reported particularly in Malaysia is increasing. According to leptospirosis data from Wahab (2015), as shown in Figure 1.2, the total number of leptospirosis cases and deaths reported over recent years has increased more than ten times, from 263 cases in 2004 to 7806 cases in 2014. Meanwhile in 2015, the number of leptospirosis cases and deaths reported has slightly decreased to 5370 cases and 30 deaths.

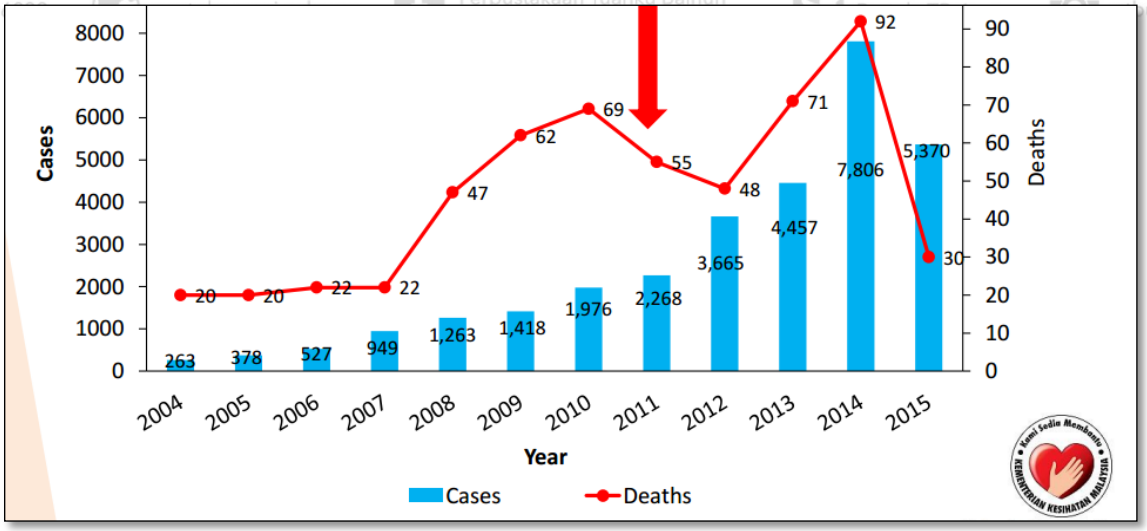

Figure 1.2. Number of leptospirosis cases and death reported from year 2004 to 2015 in Malaysia (Wahab, 2015)

According to the statistical data from Ministry of Health Hospitals in Malaysia, between 2011 and 2012, 5869 cases were reported with more than 100 fatalities. The annual incident rates of leptospirosis ranged from 7.83 to 12.49 cases

per 100 000 populations (Benacer et al., 2016). Some of the states in the West Malaysia such as Selangor, Perak, Kelantan and Pahang have the highest occurrences of incident cases. In addition, most of the cases were reported throughout heavy rainfall and flooding during the monsoon season. This critical situation should not be underestimated. Therefore, the study of leptospirosis disease specifically focusing on its geographical distributions in Malaysia is vital for the country to manage and control the occurrence of this disease.

1.3 Scope of the Research

The area of applications for the geographical distribution of infectious disease can be classified into three categories which are disease clustering, ecological analysis and disease mapping (Lawson, 2001). In disease clustering, the purpose is to observe whether a disease map reveals localized clusters of cases and to determine the location of any clusters. The aim of the analysis is to investigate potential environmental hazards. The second category is ecological analysis which focuses on the analysis of geographical distribution of the disease in relations to explanatory variable.

The third category is disease mapping which uses models to portray the disease distribution on the overall risk map. The aim of the analysis is to find the relative risk of a disease in study areas and to remove noise in the disease area. This research focuses on the third category which is the disease mapping. Lawson (2001) stated that disease mapping refers to methods that can be used to show the

geographical distribution of disease occurrences. The appearance of shaded or colored gradation scheme is included in disease mapping study to represent the incidence of disease area of interest.

1.4 Problem Statement

The focus of the disease mapping is to demonstrate the disease risk map that is estimated using relative risk estimation method by considering the disease transmission model. Generally, the disease transmission model is elaborated in the form of mathematical model to explain the route of the disease transmission. Some of the studies about leptospirosis discussed about several factors which influence the risk of human to be infected by the disease such as the transmission of leptospirosis, rainfall and flooding, climate, outdoor occupation of human and so on (see, for example, Lim et al., 2011; Lau et al., 2012; Benacer et al., 2016).

Numerous studies of leptospirosis cases have used statistical method by applying the mathematical model to study the transmission of the disease (see, for example, Triampo, Baowan, Tang, Nuttavut & Doungchawee, 2006; Pongsumpun, Manmai & Kongnuy, 2008; Pongsumpun, 2012; Pongsumpun, 2014). However, most of the previous leptospirosis studies focused on analyzing the basic reproductive number, R_0 to determine the diligence or disappearance of the disease. Furthermore, there was also a study which used multivariable logistic regression to examine the relationship between seropositivity and risk factors of leptospirosis disease (Lau et al., 2012).

Currently, the study of risk maps for leptospirosis in Malaysia is scarce. Besides, the risk map that gives detailed explanation about the situation of leptospirosis in Malaysia such as the leptospirosis risk map for children and adult humans is also not yet found. The current approaches in identifying high-low risk area of leptospirosis in Malaysia are still based on the total numbers of occurrence of leptospirosis in certain regions. A high number of leptospirosis cases in certain areas represent high risk of leptospirosis occurrence without considering other factors such as population size, age group and gender. However, it is important to consider the geographical area, number of population, age, and so on before identifying the area as high or low risk of disease occurrence. This is because the risk of disease occurrence in certain region is closely related to circumstances in geographical area such as the number of population and the age of susceptible individual in that area.

The value for relative risk needs to be estimated in order to identify the high or low risk area for disease occurrence. The application of an effective statistical method used for relative risk estimation is important to produce a good map. There are many existing methods which are used to estimate relative risk and two common methods are called Standardized Morbidity Ratio (SMR) method and Poisson-gamma model (Nor Azah & Syafiqah Husna, 2013). The SMR method basically compares the observed cases with the expected cases while Poisson-gamma model is one of the earliest examples of Bayesian mapping which uses Bayes's theorem to find the posterior expected relative risk for all regions.

Although the SMR method and Poisson-gamma model are the index to measure relative risk, they have several disadvantages. Since the SMR is a ratio, the

problems of SMR method is the value of SMR will become zero when there are no observed cases in certain regions. This relative value is not acceptable because each region should have risk to be infected with the disease. Besides, the value of SMR will become larger if the expected number of cases is smaller and the value of SMR will become smaller when the expected number of cases is larger. In the meantime, the drawback of Poisson-gamma model is this model does not allow addition of new parameter into the model.

Since the SMR method and Poisson-gamma model have some drawbacks, the SIR model and SIR-SI model were introduced to estimate the relative risk for leptospirosis in Malaysia. In order to provide more detailed information and clear interpretation for the situation of leptospirosis in Malaysia, this study introduces alternative methods to estimate the relative risk of leptospirosis for children and adult humans in Malaysia particularly by using disease transmission models for leptospirosis. This new alternative model is extended from the basic SIR-SI model for leptospirosis proposed by Triampo, Baowan, Tang, Nuttavut and Doungchwee (2006) which is discussed further in Chapter 3 as this study focuses on estimating the relative risk of leptospirosis for children and adult humans in Malaysia and the basic SIR-SI model for leptospirosis proposed by Triampo et al. (2006) does not consider the element of children and adult humans in their model.

Furthermore, this study also concentrates on significant covariates such as number of population and age of individual to help the government to control this disease from continuously spreading. Therefore, the output of this study is to produce leptospirosis maps by using the proposed method to display risks of leptospirosis in

different regions in Malaysia. This map can create awareness among the public as well as the government agencies and ultimately, this study is fundamental for leptospirosis disease prevention. Besides, the model proposed is also able to give more detailed pictures of relative risk map as it emphasises the risk of leptospirosis for children and adult humans for each state in Malaysia.

1.5 Objectives of the Research

The objectives of this research are as follows:

1. To propose an alternative method of the relative risk estimation for leptospirosis mapping based on stochastic SIR-SI Age-Structured model.

2. To estimate the relative risk for the leptospirosis disease mapping based on the discrete time and space using the stochastic SIR-SI Age-Structured model for leptospirosis transmission in Malaysia.

3. To compare the result of the relative risk estimation obtained from the SMR method, Poisson-gamma model, SIR model and stochastic SIR-SI model and stochastic SIR-SI Age-Structured model for the leptospirosis data in Malaysia by using risk map.

1.6 Significance of the Research

The increasing number of leptospirosis cases in Malaysia as shown in Figure 1.1 is alarming and requires a thorough and concerted effort for its prevention. Therefore, this research is carried out to propose a new alternative method in estimating the relative risk of leptospirosis for children and adult humans in Malaysia by overcoming the problems of Standardized Morbidity Ratio (SMR) and Poisson-gamma model used in previous studies for a better risk estimation of the leptospirosis occurrence in Malaysia. The alternative method includes the stochastic leptospirosis transmission which is suitable in representing real life situation. This stochastic transmission of leptospirosis takes into account the two groups of human age; children and adults. The new alternative method is expected to produce clearer disease maps that would give more information about the situation of leptospirosis in Malaysia.

Furthermore, the disease map is able to illustrate the high and low risk areas of leptospirosis for children and adult humans in order to identify regions that need more attention particularly from the responsible government agencies. This will lead to better ways to monitor the disease occurrence and hence developing better strategies for its prevention in Malaysia. Besides, the outcome of this study is beneficial to the communities in the study areas. The disease map can be used by the communities to identify the high risk areas of leptospirosis occurrence hence extra precautions can be taken to protect themselves against the infection.

1.7 Organization of the Thesis

There are five chapters in this thesis which have been organized to describe the study. Chapter 1 introduces the background of the research and the problem statement for this research. The objectives and significance of this research are also explained in this chapter.

In chapter 2, an explanation about leptospirosis is elaborated. The explanation includes the signs and symptoms, cause and transmission of leptospirosis, prevention and control, and treatment and vaccine development for leptospirosis. This chapter discusses previous works that have been published on modeling disease transmission especially for leptospirosis disease, and the analysis of disease mapping based on the existing method i.e. SMR method, Poisson-gamma model, SIR model and SIR-SI model. It also presents and explains the basic disease compartmental transmission model of leptospirosis. The deterministic difference equations of the SIR-SI model are also described in this chapter which later is adapted to develop a stochastic SIR-SI model.

Chapter 3 presents and explains the disease compartmental transmission model of leptospirosis by age-structured and describes the deterministic difference equations of the SIR-SI Age-Structure model. There are two types of model discussed in this chapter which are the mathematical modeling for leptospirosis among children and adult humans including the transition rate from children to adult humans and mathematical modeling for leptospirosis among children and adult humans without the transition rate from children to adult humans. This study uses the stochastic SIR-

SI Age-Structured model which omits the transition rate from children to adults in order to estimate the relative risk for leptospirosis in Malaysia.

Chapter 4 discusses the application of relative risk estimation for leptospirosis based on the existing methods and our proposed model. This chapter presents and demonstrates the application of relative risk estimation for disease mapping using observed leptospirosis data from Malaysia. Tables, graphs and maps are used to present the findings of relative risk estimation using leptospirosis data in Malaysia.

Finally, chapter 5 discusses the conclusion based on the overall results of this study. This includes the contributions of this study and some recommendations for any future studies.