# THE DEVELOPMENT OF STOCHASTIC SIR AND S($I^mI^f$)R MODELS FOR HETEROSEXUAL HIV AND AIDS DISEASE MAPPING IN MALAYSIA

SUFI HAFAWATI BINTI IDERIS

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 MASTER OF SCIENCE
(MASTER BY RESEARCH)

FACULTY OF SCIENCE AND MATHEMATICS
UNIVERSITI PENDIDIKAN SULTAN IDRIS

2016

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

## ABSTRACT

The purpose of this research is to develop an alternative method of relative risk estimation based on stochastic discrete time-space model. There are two main methods proposed which are stochastic discrete time-space SIR (Susceptible-Infected-Removed) model and $S(I^mI^f)R$ (Susceptible-[Infected male]-[infected female]-Removed) according to gender for HIV and AIDS. Firstly, the deterministic compartmental SIR model that is suitable for the transmission of HIV and AIDS is discussed. This model then is adapted to develop corresponding discrete time-space stochastic SIR model. This is followed by the discussion of the alternative method of relative risk estimation based on the stochastic SIR model. Furthermore, applications of the common methods such as SMR and Poisson-gamma using HIV and AIDS data from Malaysia is also discussed in this thesis and the results are compared in tables, graphs and maps. Subsequently, discussion on modelling of $S(I^mI^f)R$ model based on gender is discussed in the same regulation of discussion based on SIR in order to identify risk for male and female respectively. Finally, the results of analysis showed that the alternative relative risk estimation method based on the discrete time-space stochastic SIR model and $S(I^mI^f)R$ according to gender have small number of range for estimating relative risk which has been examined by using goodness-of-fit method compared to the other common approaches. As conclusion, the proposed models are better than to the common methods in estimating of relative risk. However, the SIR model can be used for the analysis of HIV and AIDS generally while $S(I^mI^f)R$ model is used to analyse relative risk specifically for gender. Besides that, the proposed model also takes into account the transmission process of the disease while allowing spatial adjustment between risks in adjacent areas. The disease risk map in this study can be used as a tool to identify states that need more attention and further action by authorities at the same time to give awareness to society about the risks in the area where they live. Furthermore, this improvement of disease mapping is able to be used as a medical or inquiry tool in any corresponding HIV and AIDS research and development study.

# PEMBANGUNAN MODEL STOKASTIK SIR DAN S($I^mI^f$)R BAGI PEMETAAN PENYAKIT HETEROSEKSUAL HIV DAN AIDS DI MALAYSIA

## ABSTRAK

Kajian ini bertujuan membangunkan kaedah alternatif bagi penganggaran risiko relatif berdasarkan model masa-ruang diskrit stokastik. Terdapat dua kaedah utama yang diperkenalkan iaitu model masa-ruang diskrit stokastik SIR (*susceptible-Infected-Removed*) dan S($I^mI^f$)R *(Susceptible-[Infected male]-[infected female]-Removed)* mengikut jantina bagi HIV dan AIDS. Pertama, suatu model kompartmen SIR bagi HIV dan AIDS dibina berdasarkan biologi jangkitan penyakit. Model ini kemudiannya disesuaikan bagi membina model diskrit masa-ruang stokastik SIR di mana bilangan jangkitan baharu bagi HIV dan AIDS adalah stokastik. Kemudian, kaedah alternatif penganggaran risiko relatif berdasarkan model SIR baharu diperkenalkan. Tambahan lagi, aplikasi kaedah biasa seperti SMR dan Poisson-gamma menggunakan HIV dan AIDS data dari Malaysia juga dibincangkan dalam kajian ini dan keputusannya dibandingkan di dalam jadual, graf dan peta. Seterusnya, model S($I^mI^f$)R berdasarkan jantina turut diambilkira mengikut aturan perbincangan yang sama seperti SIR bagi mengenal pasti risiko HIV dan AIDS untuk lelaki dan perempuan. Akhir sekali, hasil analisis menunjukkan alternatif baru anggaran risiko relatif berdasarkan diskrit masa-ruang stokastik SIR model and S($I^mI^f$)R mengikut jantina yang diperkenalkan mempunyai nilai julat g*oodness-of-fit* yang lebih kecil berbanding kaedah biasa. Kesimpulannya, model yang diperkenalkan adalah lebih sesuai bagi menganggar relatif risiko penyakit berbanding kaedah biasa. Walau bagaimanapun, model SIR adalah digunakan bagi analisis HIV dan AIDS secara amnya, sementara model S($I^mI^f$)R pula digunakan untuk menganalisis risiko relatif HIV dan AIDS spesifikasi bagi jantina. Selain itu, model baru yang diperkenalkan juga mengambilkira akan proses jangkitan penyakit sementara membenarkan pengiraan kovariasi dan kawalan ruang di antara risiko kawasan sempadan. Peta risiko penyakit yang dihasilkan dalam kajian ini dapat digunakan sebagai alat untuk mengenal pasti negeri-negeri yang memerlukan tumpuan dan tindakan yang selanjutnya oleh pihak berkuasa pada masa yang sama memberi kesedaran kepada masyarakat mengenai risiko di kawasan yang didiami. Tambahan lagi, penambahbaikan ini juga dapat digunakan sebagai medium perubatan atau penyelidikan dan pembangunan kajian HIV dan AIDS yang berkaitan.

# TABLE OF CONTENT

## LIST OF TABLE

## LIST OF FIGURE

## LIST OF ABBREVIATIONS

| | |
|---|---|
| AIDS | Acquired Immune Deficiency Syndrome |
| ART | Antiretroviral Theraphy |
| ARV | Antiretroviral Drugs |
| CAR | Conditional Autoregresive |
| CDC | Centre of Disease Control and Prevention (US) |
| DNA | Deoxyribonucleic Acid |
| GOF | Goodness-Of-Fit |
| HIV | Human Immunodefiency Virus |
| IDU | Injected Drug User(s) |
| MCMC | Markov Chain Morte Carlo |
| MOH | Ministry of Health |
| NSEP | Needle and Syringe Exchange Programme |
| OST | Opiate Substituition Theraphy |
| PLHIV | People Living With HIV |
| RNA | Ribonucleic Acid |
| SEIR | Susceptible-Expeosed-Infected-Recovered |
| SEIRS | Susceptible-Expeosed-Infected-Recovered-Susceptible |
| SI | Susceptible-Infected |
| $S(I^mI^f)R$ | Susceptible-(Infected male)-(Infected female)-Removed |
| SIR | Susceptible-Infected-Removed |

| | |
|---|---|
| SIS | Susceptible-Infected-Suscebtible |
| SMR | Standardized Morbidity Ratio |
| UNAIDS | United Programme on HIV |
| UNICEF | United Nations International Children's Emergency Fund |
| WHO | World Health Organization |

# CHAPTER 1

# INTRODUCTION



## 1.1 Research Background

Disease mapping is a prevalent method used for more than two decades to display the geographical distribution of disease occurrence (Sasivimol et al., 2004). Mapping can be defined as a collection of spatial defined objects. It is simply a display of the spatial properties of a set of objects on a map. This mapping method could be a fundamental tool to show the areas affected by the epidemic using the interpretation

of coloured maps. Disease mapping is also used to group the information in the data for all region in the study area into several group so that the region in the group has a similar risk (Nuzlinda and Abdul Aziz, 2007).

Disease mapping has two prevalent uses which are thing away noise to draw maps and as a method to assess specific hypotheses concerning incidence. Based on Olaf (2004), disease mapping, disease clustering, and geographical correlation analysis are the types of study focus that are featured in spatial epidemiology studies.

In disease mapping, the method is aimed to estimate the risk areas throughout the region where the disease map is used to highlight the areas of high or low disease occurrence. In spatial statistics, the terms of disease mapping refers to the spatial representation of morbidity and mortality rates across a region (Silva and Dean, 2012).



Meanwhile, the ecological analysis focuses on the analysis of the geographical distribution of a disease and measure the explanatory factors. This distribution is also carried out at an aggregated spatial level. Another study focus for spatial distribution is disease clustering. Cluster means a group of things of the same kind that are very close together. This spatial distribution concerns the analysis of grouping the health-related events that are temporarily related and in the proximity area.

This study will focus more on the disease mapping to estimate and present the summary in measuring the epidemiology of HIV and AIDS by showing the areas of high and low risk of the disease in Malaysia. In disease mapping, the fundamental topic to be discussed and applied is the relative risk estimation. Therefore, this research aims to introduce the alternative method of relative risk estimations based on

stochastic SIR (Susceptible-Infected-Removed), and S($I^mI^f$)R (Susceptible-[Infected male]-[infected female]-Removed) for the transmission of HIV and AIDS respectively.

Firstly, the deterministic compartmental SIR models for HIV and AIDS are discussed and represented mathematically by a system of difference equations. These models are then adapted to develop a corresponding discrete time-space stochastic SIR model where the number of newly infective HIV and AIDS are stochastic and assumed to follow a Poisson distribution. Then, an alternative method of relative risk estimation which is based on this stochastic SIR model is discussed. Furthermore, the applications of common methods such as SMR and Poisson-gamma using the data of Malaysia's HIV and AIDS are also discussed in this thesis and the results are compared in tables, graphs, and maps.



Besides, the discussion on the modelling of S($I^mI^f$)R model based on gender is also considered in this thesis. Next, the deterministic compartmental and stochastic S($I^mI^f$)R model for HIV and AIDS are discussed as the previous SIR transmission model but it focuses more on the gender whether they are male or female so it can be related to heterosexual analyses of HIV and AIDS. By performing a detailed analyses according to the gender, we could identify and estimate the relative risk for male and female respectively.

Finally, by using range as method for goodness-of-fit (GOF) these alternative relative risk estimation methods proposed which are based on the discrete time and discrete space stochastic SIR and S($I^mI^f$)R models could offer a better model for estimating the relative risk of HIV and AIDS. This is because the approach models have small number of range compared to other common approaches. These proposed

method also consider the disease transmissions as well as the spatial correlations between the neighbouring study regions.

In this analysis, discrete time-space stochastic $S(I^mI^f)R$ model according to gender is an extension for discrete time-space stochastic SIR model. Where the SIR model can be used for the analysis of HIV and AIDS generally while $S(I^mI^f)R$ model is used to analyse relative risk specifically for gender.

## 1.2 Problem Statement

In Malaysia, the relative risk estimation methods that are used to represent high and low risks are still based on the total number of cases based on behaviour of the population such as injected drug users, sex workers, men who have sex with men, and transgender. These are categorized as the population with the highest risk (Global AIDS Response Progress, 2014) instead of the total numbers of the disease occurrence across the study regions. High risk of disease occurrence is represented by the large number of cases reported without considering the geographical or spatial area of the population as well as the size or other factors that could possibly affect the disease occurrence.

For HIV and AIDS cases, most of the studies discussed on the basic reproduction number (Roxana, 2006; Basavarajaiah et al. 2012). It is very scarce to find analysis of HIV and AIDS based on the relative risk of discrete-time deterministic, discrete space deterministic, and stochastic model. Therefore, this study

would examine the relative risk estimation using discrete-time deterministic, discrete-space deterministic, and stochastic models as the new approach for HIV and AIDS cases in order to estimate the high and low risk areas for HIV and AIDS diseases in Malaysia.

### 1.3 Objectives of the Research

The objectives of this research are as follows:

1) To develop and propose a new discrete time-space stochastic model based on the difference equations in the form of SIR and $S(I^mI^f)R$ models of heterosexual infectious disease transmission for HIV and AIDS cases.

2) To propose an alternative method of relative risk estimation for heterosexual cases of HIV and AIDS based on SIR and $S(I^mI^f)R$ disease transmission models.

3) To compare the results of the relative risk estimation for HIV and AIDS based on the proposed discrete time-space stochastic SIR and $S(I^mI^f)R$ models with the SMR method and Poisson-gamma model.

4) To propose disease risk maps for HIV and AIDS in Malaysia.

## 1.4 Significant of Study

The following are the significances of the disease mapping study specifically for HIV and AIDS in Malaysia.

1. It will facilitate the authorities to identify the states that need more attention and further action to reduce the incidence of cases.

2. This risk maps could be used as a tool for public health centre in order to reveal the affected zones by the epidemic.

3. It will contribute knowledge or ideas in the research and development field as well as in the medical study of HIV and AIDS.



## 1.5 Organization of the Report

This report is divided into five chapters. Chapter 1 provides the background of the study and the problem statement. This is followed by the study objectives and the significance of the research.

Chapter 2 introduces the explanation of HIV and AIDS including the transmission phase from early stage of HIV until AIDS, biological aspects of HIV and AIDS, the diseases timeline, the signs and symptoms of the diseases, the situation of HIV and AIDS in Malaysia, and the prevention and control of the diseases. Moreover,

this topic will discuss the spatial spread of HIV and AIDS, the previous modelling disease transmission, and the analysis of disease mapping which considers the track-count data analysis. The analysis of disease mapping includes the explanation of the existing models such as Standardized Morbidity Ratios (SMR) method and Poisson-gamma model.

Next, Chapter 3 explains the methodology of the research. It begins with the discussion of the common and existing compartmental models of the diseases and followed by the compartmental and deterministic SIR model for both HIV and AIDS transmission, the stochastic SIR models for new infective HIV and AIDS, and continued with the compartmental $S(I^mI^f)R$ models for the transmission of HIV and AIDS according to gender, the stochastic $S(I^mI^f)R$ models of HIV and AIDS transmission for new infective male and female, and the last section will discuss in details about the relative risk estimation method based on the proposed stochastic models.



Chapter 4 initially presents and demonstrates the results of the relative risk estimation for disease mapping by applying the existing methods and followed by the analysis using the proposed models which are discrete time-space stochastic SIR and $S(I^mI^f)R$ models. All analyses are based on the observation of heterosexual cases of HIV and AIDS discrete-time and discrete space data in Malaysia. Furthermore, the related findings are presented and compared in tables, graphs, and maps.

Finally, Chapter 5 summarizes the findings of the thesis and discusses the conclusions, contributions, and some recommendations for future work.

# CHAPTER 2

# LITERATURE REVIEW



## 2.1 Research Background

This chapter discusses the literature related to the statistical analysis of HIV and AIDS. The first section discusses the immunology of HIV and AIDS. This is followed by the explanation of the spatial spread of the disease. Moreover, the current situation of HIV and AIDS in Malaysia and the prevention and control of the disease are

briefly presented. Finally, the modelling and the analyses of the diseases transmission and disease mapping are discussed.

## 2.2 Human Immunodeficiency Virus (HIV) and Acquired Immune Deficiency Syndrome (AIDS)

Certainly, AIDS is a disease caused by HIV. HIV is a kind of retrovirus that infects the human immune system. It makes people more vulnerable to the infection and diseases by destroying or impairing the CD4 positive T cells and macrophages which function as the key components of the cellular immune system. This virus has the potential to deplete the progression of the human immune system and cause immune deficiency. Based on Holand (2013), each day, millions of CD4 positive T cells will be produced and act as the body defenses to help fight the invading viruses and germs inside the body. CD4 positive T cell is a type of lymphocyte cell that has glycoprotein called CD4 on its surface which is very crucial for human immune system and immune response. Immune deficiency happens when the human immune system could no longer fight the infection and diseases. As a result, some of the HIV positive will develop into AIDS. Typically, the development of AIDS symptoms would finally lead to death. Until today, there is no known cure for HIV and AIDS (Ram Naresh, 2008). A person who is infected with HIV will go through various phases towards AIDS and death (Herbert and James, 1992).

HIV and AIDS are a direct transmitted diseases which are the most common form of any virus or diseases to be transmitted. Unlike indirect disease transmission, HIV and AIDS in Malaysia could transmit the virus via four major situations which are blood-to-blood or fluid, prenatal, injected drug use, and sexual contact. Direct transmitted disease occurs when there is a physical contact between a person who is infected with HIV or a person with positive HIV and a susceptible person. Figure 2.1 depicts the transmitted HIV cases reported in Malaysia since 1986 until 2012 according to mode of transmission.

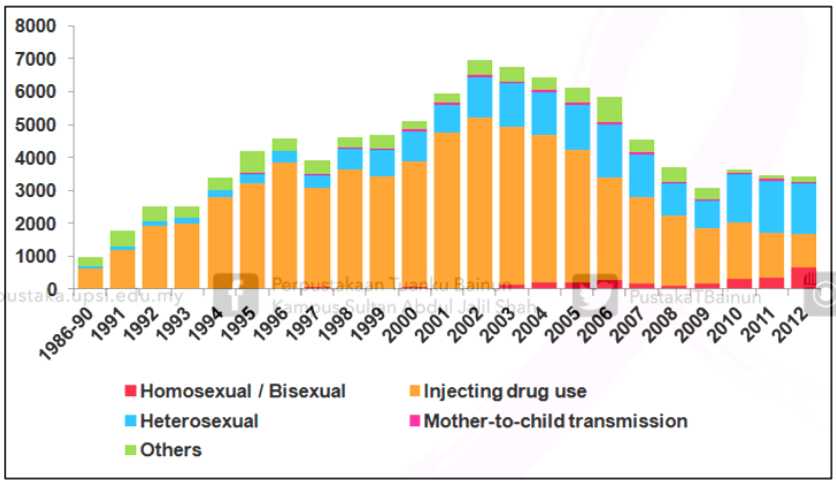


*Figure 2.1.* HIV Cases by Mode of Transmission (1986-2012). From Ministry of Health, 2014

As of the end of 2013, HIV in Malaysia is predominantly people with injected drug users as that case were the driven factor in the earlier phase of the HIV and AIDS spreading. However, the trend of infection has changed with increasingly number of infection caused by sex over the injected drug use cases started from year 2009 until 2012 as shown in Figure 2.1.

The transmission rate of HIV via sexual transmission happens when having unprotected sex including vaginal, oral or anal sex, and also by sharing sex toys which