FACTORS ASSOCIATED WITH ANTI TUBERCULOSIS THERAPY (ATT) COMPLIANCE, ATT OUTCOMES AND SURVIVAL OF PATIENTS WITH TB/HIV CO-INFECTION USING GENERALIZED STRUCTURAL EQUATION MODELING (GSEM)

RUHANA BINTI CHE YUSOF

UNIVERSITI SAINS MALAYSIA

2016

FACTORS ASSOCIATED WITH ANTI TUBERCULOSIS THERAPY (ATT) COMPLIANCE, ATT OUTCOMES AND SURVIVAL OF PATIENTS WITH TB/HIV CO-INFECTION USING GENERALIZED STRUCTURAL EQUATION MODELING (GSEM)

by

RUHANA BINTI CHE YUSOF

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

SEPTEMBER 2016

ACKNOWLEDGEMENT

The assistance and support of many individuals who have contributed to this project are gratefully acknowledged. With great pleasure, I would like to express my gratitude to my supervisor; Professor. Dr. Syed Hatim Noor and my co-supervisors; Associate Professor Dr. Aniza Abd. Aziz and Associate Professor Dr. Mohamed Rusli Abdullah. Their valuable guidance, comments, suggestions and concerns throughout this journey were really appreciated. I also like to express my sincere gratitude to my co-researchers; Dr. Anuradha P Radhakrishnan from Hospital Sungai Buloh, Selangor, Dr. Mahiran Mustafa, Dr. Anilawati Mat Jelani and Dr. Mat Zuki b. Mat Jaeb from Hospital Raja Perempuan Zainab II, Kota Bharu, Kelantan for their co-operation, advice and opinions. Not to forgot, both hospital directors for their approval to participate in this project.

I also would like to acknowledge the contributions of staff of infectious disease unit, medical specialist clinic and respiratory clinic in both hospitals for their kind assistances to me throughout the medical records reviewed and data extraction. My thanks also go to all Biostatistics and Research Methodology Unit; lecturers and staff for their contributions all along the way to complete this valuable thesis.

Last but not least, grant thanks also go to all my friends especially Anis Kausar, Adhia, Nor Hafizah, Wahidah, Nurul Jannah, Doctorate Support Group member especially Professor T Ramayyah, Dr. Othman Talib and the others who were not mention here for their assistances and supports in gathering some useful information and positive regards to get me through the entire challenging path. Finally, thanks to my parent and family in their support, understanding and belief in me to complete

this challenging journey. This study was supported by USM RU grant 1001/PPSP/812089.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADF	Asymptomatic distribution free
AFB	Acid fast bacilli
AGFI	Adjusted goodness of fit index
AIC	Akaike information criterion
AIDS	Acquired immunodeficiency syndrome
Adj OR	Adjusted Odds ratio
ART	Antiretroviral therapy
ATT	Anti tuberculosis therapy
BCG	Bacillus calmette–guérin
BIC	Bayesian information criterion
CDC	Central disease control
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CI	Confidence interval
Coef	Coefficient
Crude OR	Crude odds ratio
CRF	Case report form
df	Degree of freedom
DOTS	Directly observed therapy strategy
DNA	Deoxyribonucleic acid
ETB	Extra-pulmonary tuberculosis
FN	False negative

Freq	Frequency
GFI	Goodness of fit index
GLLAMM	Generalized linear latent and mixed model
GLS	Generalized least square
GSEM	Generalized structural modeling
HAART	Highly active antiretroviral therapy
HIS	Hospital information system
HIV	Human immunodeficiency virus
HR	Hazard ratio
HRPZ II	Hospital Raja Perempuan Zainab II
HSB	Hospital Sungai Buloh
HUSM	Hospital Universiti Sains Malaysia
ID	Infectious diseases
IRT	Item response theory
IVDU	Intravenous drug used
ll	Log likelihood
LML	Log minus log
LN	Lymph node
LR	Likelihood ratio
MC	Multicollinearity
MIMIC	Multiple indicators and multiple causes
ML	Maximum likelihood
MLMV	Maximum likelihood with missing value

MLogR	Multiple logistic regression
MOH	Ministry of Health
MOT	Mode of HIV transmission
NFI	Normed fit index
OLS	Ordinary least square
PLHIV	People living with HIV
PMEM	Preliminary main effect model
PTB	Pulmonary tuberculosis
QML	Quasi-maximum likelihood
RMSEA	Root mean square error of approximation
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
SD	Standard deviation
SE	Standard error
SEM	Structural equation modeling
SLogR	Simple logistic regression
TB	Tuberculosis
TBC	TB code
TVC	Time varying covariate
VIF	Variance inflated factor
WHO	World health organization
WLS	Weighted least square

LIST OF SYMBOLS

Δ	delta
α	alpha
β	Beta
n	number of samples
Ψ	psi
%	percent
μL	microliter
<	less than
>	more than
$\geq$	more than or equal
=	equal
$\pm$	plus minus

FAKTOR-FAKTOR BERKAITAN DENGAN TERAPI ANTI TUBERCULOSIS (ATT), HASIL RAWATAN ATT DAN KELANGSUNGAN HIDUP PESAKIT KO-INFEKSI TB/HIV MENGGUNAKAN PERMODELAN UMUM PERSAMAAN BERSTRUKTUR

ABSTRAK

Pendahuluan: Bilangan jangkitan TB/HIV yang dilaporkan di Malaysia adalah kira-kira lapan peratus daripada keseluruhan kes HIV dan kira-kira 5.9 peratus daripada jumlah kes TB yang dimaklumkan. Peratusan kes telah menurun pada setiap tahun bermula pada tahun 2007. Untuk strategi pencegahan yang lebih berkesan, model TB/HIV perlu dibangunkan terutamanya yang serasi dengan keadaan di Malaysia.

Objektif: Kajian ini telah dicadangkan untuk memodelkan jangkitan TB/HIV berdasarkan faktor-faktor yang berkaitan pematuhan terapi anti tuberculosis (ATT) dan faktor-faktor berkaitan dengan hasil rawatan TB dan juga faktor-faktor ramalan kematian dalam jangkitan TB/HIV dan juga untuk menilai model tersebut menggunakan kaedah baru yang dikenali sebagai model umum persamaan berstruktur.

Metodologi: Kajian retrospektif kohort telah mengekstrak maklumat seperti sosio-demografi, sosial dan sejarah perubatan, tanda-tanda dan gejala-gejala semasa di diagnosis, dan rawatan daripada 284 rekod perubatan dari tahun 2005 hingga 2012 di dua buah hospital kerajaan yang terpilih. Pelbagai regresi logistik telah digunakan dalam dua analisis; penentukan faktor-faktor yang berkaitan dengan pematuhan ATT

dan faktor-faktor yang berkaitan dalam hasil rawatan ATT. Sementara itu, Cox regresi digunakan dalam menentukan faktor-faktor ramalan kematian. Semua hasil telah digabungkan bersama-sama dan pembolehubah pendam diagnosis TB dimasukkan untuk menentukan model jangkitan TB/HIV menggunakan model umum persamaan berstruktur. Kesan pengantaraan juga telah dinilai dalam model tersebut.

Keputusan: Model ini telah mengenal pasti tiga faktor penting yang berkaitan pematuhan ATT (hepatitis, umur TB diagnosis dan sejarah TB sebelumnya), lima faktor penting yang berjaya dalam rawatan TB (kiraan CD4, kawasan kediaman, pematuhan ATT, tempoh ATT dan menerima HAART) dan lima faktor prognostik kematian (hasil ATT, MOT seksual, MOT IVDU, menerima HAART, tempoh ATT) termasuk satu pembolehubah yang berubah mengikutmasa (kawasan tempat tinggal). Pembolehubah terpendam diagnosis TB diukur dengan simptom batuk, demam, berpeluh malam, kehilangan berat badan dan penemuan makmal (dada x-ray, calitan sputum AFB dan pengkulturan. Gabungan semua hasil analisis secara serentak oleh kaedah baru ini telah memberi hasil yang sama seperti kaedah tradisional dengan kemajuan penambahan pembolehubah terpendam. Hasil rawatan ATT telah didapati memberi kesan pengantara pematuhan ATT kepada kematian.

Kesimpulan: Kemajuan pemodelan umum persamaan berstruktur untuk menganalisis banyak hasil keputusan taburan yang berlainan secara serentak dengan penambahan pembolehubah terpendam pada masa yang sama boleh memberi manfaat kepada ramai penyelidik untuk mengesahkan model mereka. Walau bagaimanapun, kaedah ini masih dalam pembangunan awal dan mempunyai banyak batasan yang perlu ditambah baik oleh pemaju perisian.

FACTORS ASSOCIATED WITH ANTI TUBERCULOSIS THERAPY (ATT) COMPLIANCE, ATT OUTCOMES AND SURVIVAL OF PATIENTS WITH TB/HIV CO-INFECTION USING GENERALIZED STRUCTURAL EQUATION MODELING (GSEM)

ABSTRACT

Introduction: The number of TB/HIV co-infection reported in Malaysia is about eight percent of total HIV cases and about 5.9 percent of total notified TB cases. The proportion of the cases was decreased each year started in 2007. For more effective strategies of preventive, a model of TB/HIV should be developed especially compatible with Malaysia situation.

Objectives: This study was proposed to model TB/HIV co-infection based on associated factors of anti TB treatment compliance and associated factors of TB treatment outcome and also predictor factors of mortality in TB/HIV co-infection and to assess the model using a new method known as a generalized structural equation model.

Methodology: Retrospective cohort study had extracted out information such as socio-demographic, social and medical history, signs and symptoms at diagnosis and treatment from 284 medical records from 2005 to 2012 in two selected government hospitals. Multiple logistic regression was applied in two analyses; determination of associated factors in ATT compliance and associated factors in the ATT outcome. Meanwhile, Cox regression was used in determining predictor factors of mortality. All outcomes were combined together and latent variable of TB diagnosis was added

to determine the TB/HIV co-infection model using generalized structural equation modeling. The mediating effect also was assessed in the model.

Results: The model had identified three significant associated factors of ATT compliance (hepatitis, age of diagnosis TB and history of previous TB), five significant factors of success in TB treatment (CD4 count, area of residency, ATT compliance, ATT duration and received HAART) and five significant predictor factors of mortality (ATT outcome, MOT sexual, MOT IVDU, received HAART, ATT duration) included one time varying covariate variable (area of residency). The latent TB diagnosis variable was significantly measured by symptoms of cough, fever, night sweating, loss of weight and laboratory findings (chest x-ray, sputum AFB smear and culture). The combination of all outcomes of the analyses simultaneously by a new method gave a similar result as traditional methods with advances in increments of a latent variable added. ATT outcome was suspected to mediate the effect of ATT compliance to mortality.

Conclusion: Development of generalized structural equations modeling to analyze simultaneously many outcomes of different distributions with the addition of latent variables at the same time can benefit many researchers to validate their models. However, the method was still in early development and has many limitations need to improve by the software developer.

CHAPTER 1:

INTRODUCTION

1.1 Tuberculosis

Tuberculosis (TB) is a disease that is caused by the bacterium *Mycobacterium tuberculosis* that primarily affects the lungs. It also can affect other parts of human body such as bone, brain, kidneys and spine. These are known as extra-pulmonary tuberculosis. Without any treatment, it can lead to serious health problems including death. It is spread when the air contaminated with TB bacteria from a person who has active TB disease to another person nearby who may breathe in these bacteria and become infected (CDC, 2011).

There are two conditions of tuberculosis. A condition where the infected body can fight the bacteria and stop them from growth is known as latent TB infection. At this stage, the person does not feel sick and do not have any symptoms except a positive reaction to the tuberculin skin test or special TB blood test and also cannot spread TB bacteria to the others. The second condition is when TB bacteria becomes active in the body and the immune system cannot stop them from multiplying, this is called TB disease. The person will feel sick and also can spread the bacteria to another person (CDC, 2011).

1.2 Human Immunodeficiency Virus (HIV)

Human Immunodeficiency Virus (HIV) is a retrovirus which infects humans when in contact with tissues such as those lining the vagina, anal area, mouth, eyes or through a break in the skin. HIV infection is a slowly progressive disease in which the virus is present throughout the body at all stages of the disease (Schoenfield, 2011).

There are three stages of HIV infection. The primary infection (initial stage) occurs within weeks of acquiring the virus and a symptom is characterized by a flu or mono-like illness that generally resolves within weeks. A chronic asymptomatic infection stage is a duration of infection without symptoms in an average of eight to 10 years. Lastly, the stage of symptomatic infection occurs when the defense system has been suppressed and complications have developed (AIDS stage). The symptoms are one or more unusual infections or cancers, severe loss of weight and dementia (Schoenfield, 2011).

1.3 TB/HIV co-infection

The breakdown of the body immune system is the hallmark of HIV infection. This makes HIV/AIDS patients susceptible to a variety of opportunistic infections (Walia, 2002). Tuberculosis is one of the most important and the commonest life threatening and opportunistic infections in HIV infected patients, since the pandemic of AIDS (Iyawoo, 2004; Narain & Lo, 2004; Lawn, 2005). Meanwhile, HIV is also emerging as the strongest known risk factor for the development of TB (Lawn, 2005) and the progression to active TB among those infected both with TB and HIV (Narain & Lo, 2004).

In most cases, tuberculosis infection comes first and HIV is contracted subsequently when the person achieves adolescence or adulthood. Once co-infected, the progression to active TB occurs quite rapidly (Narain & Lo, 2004).

HIV critically impairs cell-mediated host responses to *M. tuberculosis*. Numeric depletion of *M. tuberculosis*-specific CD4 lymphocytes and functional impairment of CD4 lymphocyte–macrophage interactions result in the impaired granuloma formation, ultimately leading to failure to restrict *M. tuberculosis* replication. A spectrum of histological appearances is seen; increasing immunodeficiency is associated with progressive failures of granuloma formation and increasing mycobacterial burden. The interaction between TB and HIV is bidirectional. Activation of mononuclear cells during the host response to TB leads to accelerated HIV replication, which may increase HIV load at anatomical sites involved with TB and systemically (Lawn, 2005).

Underlying HIV infection may alter the pathogenesis and clinical presentation of TB. The diagnosis of TB in patients with HIV is more difficult for the following reason as stated by; Sensitivity and reliability of tuberculin test get reduced since HIV infection causes depression of cell-mediated immunity. Only 30% to 50% of co-infected patients have a positive result. Therefore, full diagnostic evaluation should be undertaken in all patients who have clinical features compatible with TB; in HIV infected patients with pulmonary tuberculosis, sputum culture is positive for acid fast bacillus in about 90% of cases and by smear in about 50% to 70% similar to results seen in immunocompetent adults with reactivation TB and; chest X-ray abnormalities are even more non-specific in HIV infected patients than in HIV

negative patients, which may result in under diagnosis. Radiological patterns depend on the level of immunity in the host. Typical pulmonary lesions are seen only in about one-third of the HIV infected patients with clinical TB (Walia, 2002).

The aim of this study was focused on anti tuberculosis therapy (ATT) which include the factors of compliance on ATT, factors of success in ATT and survival in TB/HIV co-infection patients.

1.3.1 ATT compliance

The principles of tuberculosis treatment in HIV-infected individuals are the same as those in HIV-negative individuals which involved intensive and maintenance phases. The regimen that used in this two phases also same in both group. However, for patients who iniated anti-retrovirus therapy (ART) during TB treatment, a few factors may need to concern such as drug-drug interactions, high pill burden, overlapping drug toxicities, and immune reconstitution inflammatory syndrome (IRIS) (Bekker & Wood, 2011).

The compliance factors between this two groups may differ due to specific issues including regimen length and schedule of administration of antituberculosis drugs, timing and drug combinations of antiretroviral drugs, overlapping toxic effects, drug interactions, and occurrence of immune reconstitution in TB/HIV patients (Blanc *et. al*., 2007)

ATT compliance was defined as the extent to which the patient's behaviour matched the prescriber's recommendations (Robert Horne *et. al*, 2005). Meanwhile, non-

compliance which also well known as defaulted was defined as interruption of the treatment for two consecutive months or more (WHO, 2012).

Factors of compliance to ATT have been done in many studies in the world. Generally, the factors can be associated with the individual, disease and medication and also health services (Neves, 2010).

1.3.2 ATT outcomes

The main aim of every treatment is to cure or success in the treatment. In anti tuberculosis therapy, the outcomes were divided into two groups; successful treatment was defined as cure (negative smear of sputum in last month of treatment and on at least one previous occasion for a patient who was initially sputum smear positive) or completion of treatment (did not meet the criteria for cure or failure for a patient who completed treatment) and unsuccessful treatment was treatment default (interruption treatment for two or more consecutive months), death (died from any cause before treatment completed) or treatment failure (sputum smear remained positive at month five or later for a patient who initially sputum smear positive) (WHO, 2012).

Same as ATT compliance, there are many factors associated with the outcomes of the treatment which also can be related to individual, disease and also health services.

1.3.3 Survival in TB/HIV co-infection

Cure or success in tuberculosis therapy may increase the survival of patients. WHO (2006) has reported that more than 25% of death among TB patients is attributable to

HIV co-infection and it is the primary reason of failure to achieve TB control targets in countries with high HIV infection.

Many studies have determined that survival probability was lower in TB/HIV patients (Shaweno & Worku, 2012; Sileshi *et. al*., 2013; Roshanaei *et. al*., 2014). Low in survival probability means high in mortality. Many predictors related to the patient, disease and treatment such as compliance with the treatment (Tseng *et. al*., 2007), receiving antiretroviral treatment (Manosuthi *et. al*., 2006) and CD4 count level (Sileshi *et. al*., 2013) were associated with this event.

Identifying predictors of mortality is important to predict the prognosis of the disease in this patients and help in planning the effective interventions to reduce the mortality rate (Senbeta *et. al*., 2014),

1.4 Concepts of the Chosen Statistical analyses

1.4.1 Logistic regression

Logistic regression was used to analyze a categorical dependent variable which was measured qualitative variable like an outcome (succeed or failed), status (lived or death) or stages (mild, moderate or severe). It can be in a form of dichotomous (binary) or polytomous variable. It recommends the best fitted model or the most parsimonious model to describe the relationship between a dependent variable and a set of independent or predictor variable which represents the conditional probability of experiencing the outcome given by independent variable, $Pr(Y=1|x)$. Coefficients for its equation were computed by maximum likelihood estimates and a link function for logistic is equal to logit (Hosmer & Lemeshow, 2004):

$$logit\{Pr\ (Y = 1|x)\} = log\left\{\frac{Pr\ (Y = 1|x)}{1 - Pr(Y = 1|x)}\right\} = \beta_0 + x'\beta,$$

where β_0 is the intercept parameter and β is the vector of slope.

1.4.2 Survival analysis

Survival analysis was used to measure the time to an event of the interest. It also can handle incompletely observed time-to-event data which was known as censored data. The most common form of censoring is right-censoring which occurs when the studied event has not occurred by the time of data collection ended. There are three types of this regression; nonparametric, semiparametric and fully parametric regression which differed in their distribution assumption of failure times (Hosmer *et. al*., 2008).

Kaplan Meier was used for nonparametric maximum likelihood estimation of survival function from censored data. It is also known as a product limit estimator. Kaplan Meier estimate is a step function with discontinuities at observed event (death) times. For uncensored data, the estimation would be the empirical survival function (Kleinbaum & Klein, 2005).

Semiparametric estimation of survival analysis which used in this study was represented by Cox proportional hazard (PH) model or also known as Cox model. It is a widely used method in survival analysis. The Cox model has two component in formula; baseline hazard function ($h_0(t)$) and an exponential function of X's explanatory variables $\left(e^{\sum_{i=1}^{p} B_i X_i}\right)$. The important assumption in this model is proportional hazard must be constant over time. It this assumption is violated, the extended Cox model may required which involved time dependent variable analysis.

The formula of Cox can reduces to baseline hazard if all the X's are equal to zero or no X's in the model. That baseline hazard is unspecified. This is the reason of semiparametric model (Kleinbaum & Klein, 2005).

Meanwhile, parametric survival model is a model in which the distribution of the outcome (survival time) is specified in terms of unknown parameters, which are estimated from the data. The common distribution used in this model are the Weibull, the exponential, the log-logistic, the lognormal and the generalized gamma. This parametric model no need proportional hazard model since many of the models are acceleration failure time (AFT) models (Kleinbaum & Klein, 2005).

The Cox model is widely popular compared to parametric model because it does not rely on distributional assumptions for the outcome. Although the baseline survival function is not estimated with a Cox model, Cox-adjusted survival estimates still can produce by computer packages such as SAS, Stata, and SPSS using a complicated algorithm that generalizes the Kaplan–Meier (KM) approach. An estimation of the baseline hazard also is not necessary for the estimation of a hazard ratio since the baseline hazard cancels in the calculation (Kleinbaum & Klein, 2005).

1.4.3 Generalized Structural Equation Modeling

In the medical field, most of the recorded data in patient' private and confidential folder were in the form of statement which always contained the patient complains, laboratory investigations and evidence, the physician observations, planned of treatments, treatment effects and treatment outcomes. Even though the data was in a statement form, the majority of the data was interpreted as a categorical whether in binary, ordinal or nominal information such as signs and symptoms from the patient

complains which was interpreted as either yes/no or available/not available, based on scale for pain, classification of disease, type of treatment, type of adverse event and status of the outcome. Some of the data may also contain the information in numerical forms such as age, duration or length of treatment and value of investigation such as blood pressure, blood sugar or cholesterol. But this information, usually was interpreted as low/high or mild/severe as more meaningful interpretation of them in this field.

This valuable data were observed and combined the characteristics of diseases as a guideline for them to diagnose the disease, to plan the treatment or to estimate the outcome for the future patients. In this sense, it is important to analyze, modeling and interpret the data of patient for them to work more conveniently with the prediction model. Since every data will have more than one outcome, the data should be used many models based on the each outcome. For them to have more accurate in every decision, the models should be combined together as a disease model which contains every important aspect of the disease characters.

One of the available analyses that can fulfil the requirement is structural equation modeling. This model is built from the combination of theoretical framework and was tested to confirm the theory with the available data. It also can do the combination of many equation models simultaneously and it is suitable to build a disease model as a guideline. However, the model can be modified to fit the available data to find the best model for the management of a patient.

Structural Equation Modeling (SEM) is a comprehensive statistical approach to testing hypotheses about relations among observed and latent variable. It is combined

the measurement model and structural model into a simultaneous statistical test which is valuable in inferential data analysis and hypothesis testing. Its pattern of interrelationships directionally or non-directionally among the study constructs are specified a priori and grounded in establishing the theory (Hoe, 2008).

SEM is usually used because it permits the measurement of several variable and their inter-relationships simultaneously and also allows for simultaneous, multiple dependent relationships between variable (Hoe, 2008). The purposes of SEM are to understand the patterns of correlation or covariance among a set of variables and to explain as much of their variance as possible with the model specified (Kline, 2011).

The hypothesized causal relationships can be tested among the theoretical construct to estimate and to evaluate the structural portion of the model. The raw data are used to generate the iterations, goodness of fit indices and standardized paths (Hoe, 2008).

SEM method was originated for linear relation model (Joreskog, 1970). However, the method was modified over the time to be applied in the nonlinear structural model (Bollen, 1995). The extension of SEM made easier for non-normal data to be analyzed and interpretable.

Rabe-Hesketh (2004a) had modified a method by a combination of SEM and generalized linear mixed model to represents a generalization of multilevel regression models or generalized linear mixed models. It is known as a generalized structural equation model. The advantage of this method is it does not require the data to be balanced in any of the three criteria; no missing items in multivariate responses, the same number of units at each level in multilevel design and balanced

covariates, which need the same sets of values for each higher-level unit in a balanced multilevel design.

Before that, Muthen (1979) had developed a structural equation model that involved dichotomous responses latent variable which allowed the representation of the causal relations between responses and exogenous variable. Then, the analysis became more widely used as compatible with different types of data, such as continuous, ordinal or multinomial data with latent variable. His team also developed a software namely MPlus to run the SEM analysis and deal with this type of data.

The choice of estimation in SEM has played an important role to get a correctly specified model. An estimation procedure such as maximum likelihood (ML), generalized least square (GLS) and weighted least squares (WLS) mostly will depend on a distribution of a data. Meanwhile, Bayesian estimation will be the chosen for a data with categorical responses variable.

This new method recently was applied in a few studies. Musenge *et al*. (2013) in a study of understanding TB/HIV mortality in children had used a structural equation model to modeling a complex relationship of multiple exposures and the mortality of TB/HIV in children as an outcome.

This approach whether structural equation modeling or its extension was also used in the estimation of the direct, indirect and total effects. For example, Karpa *et al*. (2009) had used structural equation modeling and Cox regression to determine the associated factors of visual impairment to mortality in older people and assessed the direct and indirect effects that link between the pathway. Meanwhile, Christ *et al*.

(2008) had used a generalized linear structural model to estimate the direct effect of visual impairment on health, disability and mortality and also estimated the indirect effects of visual impairment on mortality through health and disability mediators.

In time to time, many researchers have modified and developed new estimation methods to compatible with different types of data. Each extension of SEM has advantages and disadvantages. They have their owned limitations. This same goes into software development. Some software can analyze different types of data and some may be cannot do that. The common software to analyze SEM are AMOS, MPlus, SAS, LISREL, EQS, STATA, R and SmartPLS.

1.5 Problem Statement and Research Gap

Medicine is a wide field to perform applications of statistics in research. Many analyses can be applied in this field. One of the new methods in statistics is a structural equation modeling. Applications of structural equation modeling in medicine are in a development phase. Majority of them were in forms of continuous covariates and were analyzed using a linear structural equation modeling. However, since much of information in medical data were in forms of categorical data such as signs and symptoms (Yes/No) or more meaningful if the variables were in categorical forms. For example, age was more meaningful if it is interpreted as young and old or grouping them as infants, toddlers, teenagers and adults. A classical approach of structural equation model has to be modified to be compatible with the data without any problem in the validity of the analysis.

Generalized structural equation modeling is one of the options that we can analyze this type of data. But its application especially in medicine is still new. From literature search, only one study had applied this method to an ophthalmology study using MPlus software (Christ *et. al*., 2008). So this study would like to apply this method using a new program in STATA software in TB/HIV co-infection model specified more on multiple regression equations involving simultaneous equations and latent variables with aiming it will contribute knowledge to other researchers to apply this method in their research.

As a conclusion, TB/HIV co-infection is a complex disease which demands the application of an advanced analysis that able to link multiple causations and outcome clearly using a path diagram i.e GSEM. At present, not many studies had applied the analysis of GSEM in complicated medical data such in this disease. Hopefully, this study would contribute as a guideline for other medical researchers to apply this method in their research.

CHAPTER 2

LITERATURE REVIEW

2.1 HIV Case Definition

WHO (2007) had published a guideline for HIV case definition. A case of HIV infection is defined as an individual with HIV infection irrespective of clinical stage (including severe or stage 4 clinical disease, also known as AIDS) confirmed by laboratory criteria according to country definitions and requirements. WHO provides a simplified HIV case definition designed for reporting surveillances. However, countries should develop and review their testing algorithms regularly for diagnosis and surveillance purposes.

HIV infection is diagnosed based on positive HIV antibody testing (rapid or laboratory-based enzyme immunoassay) and confirmed by a second virological test obtained from a separate determination. First, clinical criteria for the diagnosis of advanced HIV in adults and children with confirmed HIV infection or presumptive or definitive diagnosis of any stage 3 or stage 4 conditions. Second, immunological criteria for diagnosing advanced HIV in adults and children five years or older with confirmed HIV infection, which is CD4 count less than 350 per mm^3 of blood in an HIV-infected adult or child (WHO, 2007).

2.2 Tuberculosis Case Definition

MOH (2002) had a guideline for tuberculosis case definition:

1. Pulmonary tuberculosis: Tuberculosis involving the lung parenchyma.
2. Pulmonary tuberculosis, smear positive:
    i. Tuberculosis in a patient with at least two initial sputum smear examinations (direct smear microscopy positive for acid fast bacilli (AFB).
    ii. Tuberculosis in a patient with one sputum smear examination positive for AFB and radiographic abnormalities consistent with active pulmonary tuberculosis as determined by the treating doctor.
    iii. Tuberculosis in a patient with at least one sputum smear examination positive for AFB and sputum culture positive for *M. tuberculosis*
3. Pulmonary tuberculosis, smear negative:
    i. Tuberculosis in a patient with at least three sputum smear examinations negative for AFB and with radiographic abnormalities consistent with pulmonary tuberculosis, determined by a doctor followed by a decision to treat the patient with a full course of anti-tuberculosis therapy.
    ii. Tuberculosis in a patient whose initial sputum smears were negative, who had sputum sent for culture initially, and whose subsequence sputum culture result is positive for *M. tuberculosis*
4. Extrapulmonary tuberculosis: Tuberculosis of organs other than the lung parenchyma. Diagnosis should be based on at least one culture-positive specimen from an extrapulmonary site or histological or strong clinical

evidence consistent with active extrapulmonary tuberculosis followed by a decision by a doctor to treat with a full course of anti-tuberculosis therapy.

5. Pulmonary with extrapulmonary tuberculosis: Tuberculosis involving the lung parenchyma as well as any other part of the body.

2.2.1 Definition of Terms in Tuberculosis Case

1. New case: A patient who has never had treatment for tuberculosis or has taken anti tuberculosis drugs for less than 4 weeks duration in the past.
2. Relapse case was divided into two categories:
    a. Sputum positive relapse: A patient who has been declared cured of any form of tuberculosis in the past by a doctor after one full course of chemotherapy and has become sputum smear positive.
    b. Sputum negative relapse: A patient who has been declared cured of any form of tuberculosis in the past by a doctor after one full course of chemotherapy and has developed active disease based on bacteriological, histological or clinical and radiological assessment
3. Chronic case: A patient who remained or becomes smear positive again after completing a fully supervised retreatment regimen.
4. Treatment failure: A patient who while on treatment, remained or become again smear positive 5 months or later after commencing treatment. It is also a patient who was initially smear negative before starting treatment and become smear positive after the second month of treatment.
5. Treatment after interruption: A patient who interrupts anti tuberculosis treatment for 2 months or more, and then returns to the health service with

smear positive sputum. Sometimes smear negative, but still with active tuberculosis as judged on clinical and radiological assessment.

6. Transferred in/out case: A patient transferred from/to another centre for continuation of treatment of tuberculosis. A transfer implies that the centre to which the patient is transferred undertakes the responsibility of continuing to treat the patient and supervising the progress. A patient is not considered to have been transferred if he/she presents at another treatment centre merely to obtain treatment.

2.3 Epidemiology of TB/HIV Co-infection

2.3.1 Global Situation

The number of people living with HIV (PLHIV) has stated to be increased from year to year. In 2009, the number of PLHIV was 33.3 million (31.4 million-35.3 million). From there, 30.8 million (29.2 million-32.6 million) were adult, 2.5 million (1.6 million- 3.2 million) were children below 15 years old and 15.9 million (14.8 million- 17.2 million) were women (WHO, 2009)

Data for people newly infected with HIV in 2009 was shown that the total was 2.6 million (2.3 million- 2.8 million) with 2.2 million (2.0 million-2.4 million) were adults and 370,000 (230,000 – 510,000) were children below 15 years old (WHO, 2009). WHO has reported that Sub-Saharan Africa was 80% of the global burden of HIV associated TB in 2009. The magnitude of HIV infection is known to be greatest in sub-Saharan African, where as many as a third of all patients with active tuberculosis are HIV infected.

However, in other regions of the world, where the overall prevalence of tuberculosis is not as high as in developing countries, the interaction between the two pathogens may also be a substantial problem. In regions such as the Western Pacific and Southeast Asia, WHO data indicated that TB/HIV co-infection is much less common than in Africa, although it is likely that the estimates are on the lower side because of under reporting of HIV infection (Walia, 2002).

The seroprevalence of HIV infection that was studied by Kassu *et al.* (2007) in northwest Ethiopia was 52.1% from the total of 257 TB patients with pulmonary TB and extrapulmonary TB were diagnosed in 64.2% and 35.8% of the patients, respectively. The study also concluded that the co-infection with HIV was high in patients with TB.

2.3.2 Situation in Asia

Asia has the largest numbers of tuberculosis cases (60% of the global total) and 8.55 of 39.50 million in HIV seroprevalence (Vermund & Yamamoto, 2007). It also reported, that 40% to 70% of HIV patients had tuberculosis infection and 40% of persons were dying with HIV were attributable to tuberculosis co-infection.

In mainland China, a cross sectional study was conducted through a standardized questionnaire was done by Jiang *et al*. (2008). The three year survey study from January 2003 to December 2005 which involved 241 eligible patients with TB/HIV co-infection was shown that 66.0% of the patients were co-infected with pulmonary TB and 14.5% were co-infected with haematogenous disseminated pulmonary TB.

Meanwhile, 40.7% of the total numbers of cases in the study were co-infected with extra-pulmonary TB.

2.3.3 Situation in Malaysia

In Malaysia, WHO (2011) has reported that in 2009, there were 18,102 cases of TB with 1,582 of death cases and the incidence rate was 63.95. The new cases were 16,921 and 1,181 were retreatment cases.

Data from Ministry of Health (MOH, 2012) has shown that in 2012, there were 1347 cases of TB/HIV co-infection (5.9%) with contribution from Selangor 192 cases (14.3%) and Kelantan 179 cases (13.3%). Both states were the highest and the second highest of TB/HIV co-infection in Malaysia in 2012. In 2009, from 15,192 (84%) TB patients who were screening for HIV, there were 1,644 TB patients who were infected with HIV (WHO, 2011) and the number of the cases was reduced to 1629 in 2011 (MOH, 2012). The data showed that the number of cases was decreased year to year starting from 2007 (Figure 2.1).

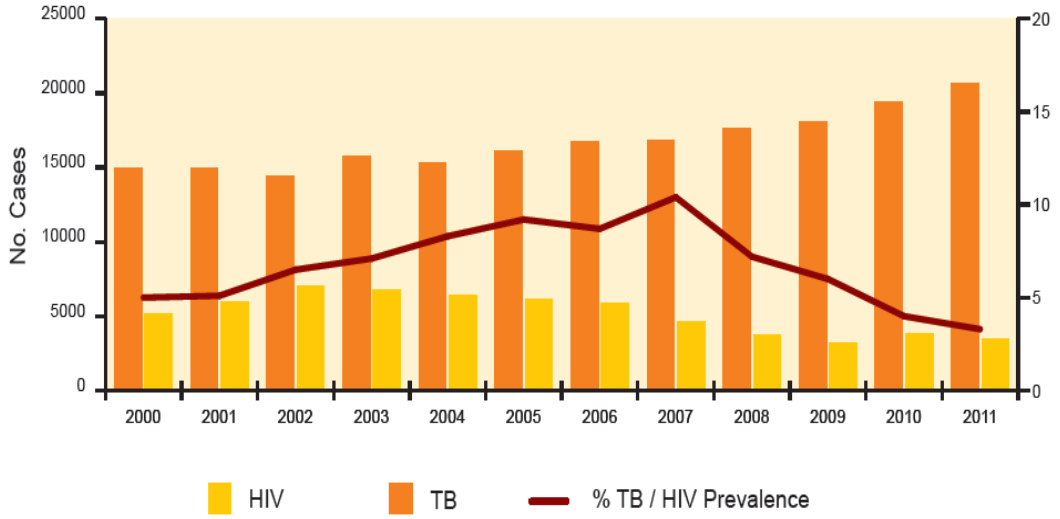

Figure 2.1: TB/HIV co-infection from 2000 to 2011 (adapted from MOH, 2012)

A retrospective record review study done by Nissapatorn *et al*. (2005) at National Tuberculosis Centre in Malaysia in order to compare the characteristics of HIV positive patients with pulmonary and extra-pulmonary tuberculosis was found that from a total of 252 HIV positive patients, the incidence of pulmonary tuberculosis (PTB) was 78.6% with 10.6% from the group was disseminated tuberculosis. Meanwhile, the incidence of extra-pulmonary tuberculosis (ETB) was 21.4%.

In most cases, the clinical presentation of tuberculosis in patients with HIV is indistinguishable from those patients who do not have HIV infection. However, some HIV positive TB patients, particularly at an advanced stage of HIV may present with an atypical pattern, with a higher proportion of cases tending to have a negative sputum smear. Nonetheless, sputum smear examination remains an essential component in TB diagnosis, even in countries where HIV infection is common, because of its ability to identify infectious cases (Narain & Lo, 2004).

From the study by Jiang *et al*. (2008), the clinical manifestations in the majority of the patients were fever (87.6%), fatigue (61.0%) and weight loss (60.2%).

Other studies conducted by Nissapatorn *et al*. (2005) and Kassu *et al*. (2007) also stated that the main clinical presentations for PTB were cough, loss of appetite or / and loss of weight and fever. There were also significant associations between cough, sputum and haemoptysis (Nissapatorn *et al*., 2003). But when they control the confounder by Multiple Logistic regression, the only significant associations were cough and haemoptysis. Most of the PTB patients were came with opacities with or without cavitation of abnormal radiological findings. Meanwhile the common site for

ETB was miliary TB with the risk factors for this type of TB were men, intravascular drug users and specific racial origin (Nissapatorn *et al*., 2005).

The study also found that lower CD4 count was more potential to develop ETB rather than PTB. The duration of treatment to success was longer in ETB (nine months) compared to PTB (six months).

The risk factors that associated with TB diagnosis were younger age, lower recent CD4 T cells count, duration of antiretroviral therapy and living in high TB burden countries. Patients aged more than 40 years old had significantly lower rate of TB diagnosis than that of patients aged 30 or younger (adjusted Hazard ratio (HR)=0.47;95%CI: 0.28, 0.79). A higher CD4 T cells count was associated with a significantly lower rate of TB diagnosis. Compared to patients who were not receiving antiretroviral therapy, there was a significant increase in the rate of TB diagnosis within 90 days after initiating therapy (adjusted HR=2.52, 95% CI: 1.31, 4.84). The rate of TB diagnosis was significantly lower among patients living in countries with low/intermediate TB burden (adjusted HR=0.28, 95% CI: 0.17, 0.45) (Zhou *et al*., 2009).

2.4 Associated Factors of ATT Compliance

ATT compliance or adherence to the treatment may be influenced by many factors. This situation can be particularly challenging since the duration of treatment take a long period (usually six months or longer), requirement of combination therapy, and side effects may be unpleasant. Furthermore, when patients experience rapid

improvement in symptoms, the decision to discontinue the treatment may take a place even though the treatment is not complete yet (Reichman & Lardizabal, 2013).

Some factors such as the distance to health facilities (Castelnuovo, 2010) which consumes time , transportation, financial status (Neves, 2010) and social support also influence patients to comply with the treatment.

Besides that, the support and counseling session from the health care staff may motivate patients to adherence to the treatment. The staff also can provide the knowledge of the disease and the important of the treatment to the patients (Neves, 2010).

Directly observed therapy is a strategy to improve the treatment adherence. However, from the existing trials, DOTS did not provide a solution to poor adherence in TB treatment. Given the large resource and cost implications of DOTS, policy makers might want to reconsider the strategy that depends on direct observation (Karumbi & Garner, 2015)

Anaam *et. al.* (2013) had suggested that reducing travelling and waiting times for TB patients may improve compliance rates by expansion of directly observed treatment short-course near to patients' homes and involving additional staff.

2.5 Associated Factors of ATT Outcomes

Compliance to a tuberculosis treatment may give a positive impact on anti tuberculosis treatment success. Most of the studies divided the treatment into two categories; success and unsuccess. Patients who cured or completed the treatment

was grouped into success group. Meanwhile, for patients who failed in treatment, defaulted, transfer or died was located in unsuccess group (Okanurak *et. al*., 2008; Shaffer *et. al*.,2012; Ismail & Bulgiba, 2013a)

Ismail & Bulgiba (2013a) study which was claimed as the first report in Malaysia on identifying the risk factors of unsuccessful TB treatment outcome in HIV infected patient had identified four associated factors of unsuccessful in TB treatment after twelve months initiated the TB treatment among TB/HIV co-infected patients. The factors were intravenous drug users (IVDU), not receiving anti-retroviral therapy, lymphadenopathy and low in serum albumin.

A cohort study which was conducted at Bangkok, Thailand had predicted the factors associated with successful in TB treatment. The predictor factors were female, monthly wage income patients, patients with moderate or high levels of knowledge in TB and its treatment and patients with adverse effects (Okanurak *et. al*., 2008).

In South India, Vijay *et. al.* (2011) had determined the factors associated with unsuccessful in TB treatment in TB/HIV co-infected patients such as pulmonary TB, retreatment cases and non initiation of antiretroviral therapy (ART). Meanwhile in North India, the factors of unsuccessful in TB treatment were retreatment cases and CD4 cell count less than 200cell/µL (Sharma *et. al.,* 2014).

2.6 Survival and Predictors Factors of mortality in TB/HIV co-infection

A negative TB tuberculin test and low CD4 count were associated with an increase in mortality, implying that the extent to which immune function is suppressed is of the

utmost importance in the prediction of the survival rate of patients with TB/HIV co-infection (Serrat *et al*., 1998; Jiang *et al*., 2008).

Survival in a similar group of those who never had TB but were matched for $CD4^+$ T-cell count showed that at each level of $CD4^+$ T-cell count, survival was worse for patients who had TB. It appears that TB accelerates the natural history of HIV infection and leads to earlier death.

Klautau & Kuschnaroff (2005) were observed lower $CD4^+$ T-cell mean values in cases of treatment failure and death. There was a significant correlation between the $CD4^+$ T-cell values and the TB outcome at six time points of the study (0 months, 2 months, 4 months, 6 months, 10 months and 15 months). Meanwhile there was a significant correlation between $CD8^+$ T-cell values and TB outcome at the first and third assessment point of the study. A correlation between HIV viral load values and the response to the treatment was also observed at the end of the treatment. The study also found that there was a possible relationship between TB outcome and laboratory parameters, lower mean values of haemoglobin, haematocrit, platelet and leucocytes among the cases of treatment failure. As a conclusion, Pancytopenia and low levels of $CD4^+$ T-cell and $CD8^+$ T-cell at the initial time point of the study were correlated with an unfavourable outcome of TB and can be considered as a potential predictor factors.

Tseng *et al*. (2009) done a study on "effect of free treatment and surveillance on HIV- infected persons who have tuberculosis, Taiwan, 1993-2006" has reported that the surveillance reporting and management and with the availability of free HAART has increased the survival rates of persons co-infected with HIV and TB.