Preliminary Estimation of Risk Factors That Associated With The Prevalence of Tuberculosis

Wan Muhamad Amir W Ahmad
Jabatan Matematik, Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu (UMT)
21030 Kuala Terengganu,
Terengganu Malaysia

Nor Azlida Aleng
Jabatan Matematik, Fakulti Sains dan Teknologi,
Universiti Malaysia Terengganu (UMT),
21030 Kuala Terengganu,
Terengganu Malaysia

Norizan Mohamed
Jabatan Matematik, Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu (UMT),
21030 Kuala Terengganu,
Terengganu Malaysia

Zalila Ali
Pusat Pengajian Sains Matematik,
Universiti Sains Malaysia (USM)
11800 Minden Pulau Pinang, Malaysia.

Abstract - Disseminated tuberculosis (TB) is a contagious bacterial infection that has spread from the lungs to other parts of the body through the blood or lymph system. TB remains one of the world’s leading infectious causes of death among adults. TB most commonly affects the lungs but also can involve almost any organ of the body [2]. Because of that, we conduct this study in order to better understand evaluate the latent form of TB infection and their relation to the other disease. We used Risk Estimation and MLR methods to gain the factors that contribute to the spread to the other disease. Dataset consists of 284 Miliary TB patients and they were collected directly from Medical Unit Records Department. These findings are concern for public health authorities, and reveal tuberculosis itself as a potential factor for other diseases.

Key words - Mycobacterium; Tuberculosis; Risk Factors

I. INTRODUCTION
Tuberculosis or tubercles bacillus is a common and often deadly infectious disease which is caused by various strains of Mycobacterium tuberculosis in humans. This bacterium usually attacks lungs, heart and other parts of the body. TB is spread through the air from one person to another (spread through the air with coughing or sneezing). Tuberculosis (TB) and human immunodeficiency virus (HIV) infections are the deadliest chronic infections globally [8]. Although each is deadly alone, they are deadlier together, with TB causing one-quarter of AIDS-related deaths and HIV infecting at least 15% of patients with TB worldwide [2,3].

II. CAUSES AND RISK FACTORS
Tuberculosis (TB) infection can develop after inhaling droplets sprayed into the air from a cough or sneeze by someone infected with the Mycobacterium tuberculosis bacteria. Small areas of infection, called granulomas (granular tumors), develop in the lungs. The usual site of TB is the lungs, but other organs can be involved. In the U.S., most people with primary tuberculosis get better and have no further evidence of disease. Disseminated TB develops in the small number of infected people whose immune systems do not successfully contain the primary infection [2]. Disseminated disease can occur within weeks of the primary infection, or may
lie inactive for years before causing illness. Infants, the elderly, those infected with HIV and those who take immune-suppressing medications are at higher risk for disseminated TB, because of their weaker immune systems [2, 4, 5]. In disseminated disease, organs and tissues affected can include:

- Bones and joints
- Larynx (voice box)
- Eye
- Stomach
- Lymph nodes
- Urinary system
- Organs of the male or female urinary and reproductive systems
- Bronchus
- Skin
- Intestines
- Lining of the heart (pericardium)
- Lining of the abdominal cavity (peritoneum)
- Lining of the brain and spinal cord (meninges)

II. DATA AND METHODS

Study Population

We studied patients with TB infection. A total of 284 eligible patients were selected. They were diagnosed to have TB based on WHO criteria.

Determine the associated factors using MLR

The aim of this study is to determine the prevalence of TB and explore the risk factors which are potential to get infected by TB patients in Malaysia. This statistical method is very useful for predicting the independent variable which is contributing to the Miliary TB. It reduces many variables to a few factors. It also produces several linear combinations of observed variables which are called as factors. The factors summarize the pattern of correlations in the observed data. Because there are normally fewer factors than observed variables and because factor scores are nearly uncorrelated, use of factor scores in other analyses may be very helpful [9]. The multiple linear regression model assumes a linear (in parameters) relationship between a dependent variable \( y_i \) and a set of explanatory variables \( x'_i = (x_{i0}, x_{i1}, ..., x_{ik}) \). \( x_{i0} \) is also called an independent variable, a covariate or a regressor. The first regressors \( i = 1 \) is a constant unless otherwise specified.

Logistic Regression Analysis

Results from path analysis shows that there is an association between tuberculosis of other organs with diseases of the nervous system and infection with parasite diseases. To get the statistical association, we perform the logistic regression [10]. Let us define following dichotomous variables as follows:

- \( Y_{ij} = 0 \), if no tuberculosis of other organs and \( Y_{ij} = 1 \), if tuberculosis of other organs

Then let us define the following model 1a:

\[
g(X_{ij}) = \beta_{0j} + \beta_{1j} \text{Infect} + \beta_{2j} \text{Nervous}
\]

Determine the Risk Estimate Using Chi-Square Method

For 2×2 tables, a measure of the strength of the association between the presence of a factor and the occurrence of an event. If the confidence interval for the statistic includes a value of 1, we cannot assume that the factor is associated with the event. The odds ratio can be used as an estimate or relative risk when the occurrence of the factor is rare (Amir & Mustafa, 2009). We can define the relative risk of disease as:

\[
p_1 = P(\text{Disease} | \text{factor present}) = P(D | F)
\]

\[
p_2 = P(\text{Disease} | \text{factor absent}) = P(D | \overline{F})
\]

The relative risk as

\[
RR = \left( \frac{p_1}{p_2} \right) = \frac{P(D | F)}{P(D | \overline{F})}
\]

Path Analysis
Path Tracing Rules

In 1921, Wright has proposed a simple set of path tracing rules, for calculating the correlation between two variables [8]. The rules for path tracing are:

i. We can trace backward up an arrow and then forward along the next, or forwards from one variable to the other, but never forward and then back.

ii. We can pass through each variable only once in a given chain of paths.

iii. No more than one bi-directional arrow can be included each path-chain.

Sample Size required

The calculation of sample size was done using Power and Sample Size Calculation (PS) software with the significance level ($\alpha$) 0.05 and the power of study (1-$\beta$) of 80.0%, see [1]. Parameter involved:

i. Type I Error = 5.0%

ii. Power = 80.0%

iii. $M = 1$

iv. $P_0$ = Based on literature review

v. $P_1$ = Based on Expert Opinion

<table>
<thead>
<tr>
<th>No.</th>
<th>Variables</th>
<th>$P_1$</th>
<th>$P_0$</th>
<th>M</th>
<th>Type I error</th>
<th>Power</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CD4 Cell count (Nissapatorn. et al., 2003)</td>
<td>0.52</td>
<td>0.36</td>
<td>1</td>
<td>5%</td>
<td>80%</td>
<td>150 patients</td>
</tr>
<tr>
<td>2</td>
<td>Heterosexual (Nissapatorn. et al., 2003)</td>
<td>0.44</td>
<td>0.28</td>
<td>1</td>
<td>5%</td>
<td>80%</td>
<td>140 patients</td>
</tr>
<tr>
<td>3</td>
<td>Intravenous drug user (IVDU) (Nissapatorn. et al., 2003)</td>
<td>0.56</td>
<td>0.4</td>
<td>1</td>
<td>5%</td>
<td>80%</td>
<td>152 patients</td>
</tr>
<tr>
<td>4</td>
<td>Blood transfusion (Nissapatorn. et al., 2003)</td>
<td>0.41</td>
<td>0.25</td>
<td>1</td>
<td>5%</td>
<td>80%</td>
<td>134 patients</td>
</tr>
<tr>
<td>5</td>
<td>Age (Nissapatorn. et al., 2003)</td>
<td>0.53</td>
<td>0.37</td>
<td>1</td>
<td>5%</td>
<td>80%</td>
<td>151 patients</td>
</tr>
</tbody>
</table>

*Determined by expert opinion

According to the Table 1, the minimum requirement sample size need is 152 observations. In this analysis, we already collect data from 284 patients. A standard form was used to review medical records and to collect data on each patient.
TABLE II. Explanation of the Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Explanation of the Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age of Patients</td>
</tr>
<tr>
<td>Gender</td>
<td>Patient’s Gender</td>
</tr>
<tr>
<td>Married_Status</td>
<td>Married Status (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Tuber_Organ</td>
<td>Tuberculosis of other organs (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Infect</td>
<td>Infection with parasite diseases (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Endo</td>
<td>Endocrine, Nutritional And Metabolic Diseases</td>
</tr>
<tr>
<td>(0 = No and 1 = Yes)</td>
<td></td>
</tr>
<tr>
<td>Nervous</td>
<td>Diseases of The Nervous System (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Hypertensive</td>
<td>Hypertensive Disease (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Respiratory</td>
<td>Influenza, Pneumonia and other Acute Lower Respiratory Infections</td>
</tr>
<tr>
<td>(0 = No and 1 = Yes)</td>
<td></td>
</tr>
<tr>
<td>Infections</td>
<td></td>
</tr>
<tr>
<td>Oesophagus</td>
<td>Diseases of Esophagus, Stomach and Duodenums, Stomach and</td>
</tr>
<tr>
<td>(0 = No and 1 = Yes)</td>
<td>Duodenum</td>
</tr>
<tr>
<td>Dorsoptathies</td>
<td>Spondylarthropathies (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Diseases_Genital</td>
<td>Having Diseases of The Genito-Urinary System (0 = No and 1 = Yes)</td>
</tr>
<tr>
<td>Descendent</td>
<td>Descendant</td>
</tr>
<tr>
<td>(1=Malay, 2=Chinese, 3= Indian &amp; 4=Others)</td>
<td></td>
</tr>
</tbody>
</table>

III. RESULTS

MLR Analysis

Data entry and analysis was done using statistical Package for Social Sciences (SPSS) version 19. The association between Age and all the variables were examined by MLR. All the p values were significant thus the model had fit. The main effect of the model was also checked for interaction by using two way interaction tests and if this was not significant, there were no significant interaction between each variable in the final model.

In this analysis, we use MLR approach to determine which factors are strongly associated with Miliary TB diseases. We summarized the significant variables in Table 3. These variables had a direct relationship with age.

TABLE III. Estimates of Parameters of MLR Model

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>27.300</td>
<td>2.618</td>
<td>10.427</td>
</tr>
<tr>
<td>Married Status</td>
<td>22.744</td>
<td>1.806</td>
<td>12.593</td>
</tr>
<tr>
<td>Descendant</td>
<td>-2.226</td>
<td>0.637</td>
<td>-3.493</td>
</tr>
<tr>
<td>Gender</td>
<td>4.614</td>
<td>1.881</td>
<td>2.452</td>
</tr>
<tr>
<td>Diseases of The Genito-Urinary System</td>
<td>21.586</td>
<td>10.652</td>
<td>2.027</td>
</tr>
</tbody>
</table>

Dependent Variable: AGE

The results in Table 3 suggested that there is a direct positive and negative relationship between Age with 4 important variables. Results shown in Table 3 illustrated that four variables were include in the equation; Married Status ($\beta = 22.744, p < 0.000$), Descendant ($\beta = -2.226, p < 0.000$) and Gender ($\beta = -0.305 p < 0.015$) and Diseases of The Genito-Urinary System ($\beta = 21.586, p < 0.044$) were found to have a direct and positive effect on Age.

Chi-Square Analysis

Hypothesis

$H_0 :$ Parasite infection among TB patients and marital status is not related to each other.

$H_1 :$ Parasite infection among TB patients and marital status is related to each other.

TABLE IV. Chi-Square Test for Association

<table>
<thead>
<tr>
<th>Married Status</th>
<th>0</th>
<th>1</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infectious with Parasite Diseases</td>
<td>109 (40.5%)</td>
<td>160 (59.5%)</td>
<td>0.046</td>
</tr>
<tr>
<td>Yes</td>
<td>10 (66.7%)</td>
<td>5 (33.3%)</td>
<td></td>
</tr>
</tbody>
</table>

Parasite infection among TB patients and marital status is related to each other.

ISSN : 2249-913X Vol. 2 No. 1 Mar-May 2012 9
Interpretation and Results Presentation

Because \( p = 0.046 \) is less than 0.05, we reject null hypothesis and we can conclude that there is significant association between parasite infection with marital status.

Risk Estimation for Parasite Infection with Marital Status

We have checked all the possible ways of risk estimates. However, all the 95% confidence interval includes 1; these indicate that an invalid odds ratio, except for risk estimation between married status with parasite infection. We summarize the finding in Table 4.

<table>
<thead>
<tr>
<th>Value</th>
<th>Odds Ratio For Married Status (0 / 1)</th>
<th>For Cohort Infectious And Parasitic Diseases = No</th>
<th>For Cohort Infectious And Parasitic Diseases = Yes</th>
<th>N of Valid Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.341</td>
<td>0.945</td>
<td>2.773</td>
<td>284</td>
</tr>
<tr>
<td>95% Confidence</td>
<td>0.113</td>
<td>0.889</td>
<td>0.973</td>
<td></td>
</tr>
<tr>
<td>Interval</td>
<td>1.024</td>
<td>1.004</td>
<td>7.903</td>
<td></td>
</tr>
</tbody>
</table>

Table 4 show the odds of parasite infection among married patients are 2.7 times compared to the non-married patients.

Path model Analysis

Direct and Indirect Effects

Figure 2 summarized the reduced path analysis for factors of Tuberculosis of other organs. In reduced model, some of the path deleted, because they are not contributed to the model. Because of that, the whole path shown in the Figure 2 is significant.

For direct effect, we would note that:
Diseases of the nervous system and infection with parasite diseases have direct effect upon Tuberculosis of other organs.

For indirect effect, we would note that:

Fig. 2. Reduced path analysis for factors of Tuberculosis of other organs
1. Diseases of esophagus, stomach and duodenums, has no direct effect upon Tuberculosis of other organs but have an indirect through diseases of the nervous system.
2. Hypertensive disease has no direct effect upon Tuberculosis of other organs but have an indirect through diseases of the nervous system.
3. Endocrine, nutritional and metabolic diseases have no direct effect upon tuberculosis of other organs but have an indirect through infection with parasite diseases.
4. Endocrine, nutritional and metabolic diseases has no direct effect upon tuberculosis of other organs but have an indirect through diseases of the nervous system.
5. Marital status has no direct effect upon Tuberculosis of other organs but have an indirect through infection with parasite diseases.
6. Respiratory infection has no direct effect upon tuberculosis of other organs but have an indirect through infection with parasite diseases.

Logistic Regression Analysis

Results from path analysis shows that there is an association between tuberculosis of other organs with diseases of the nervous system and infection with parasite diseases. To get the statistically association, we perform the logistic regression. Let us define following dichotomous variables as follows:

\[ Y_{ij} = 0, \text{ if no tuberculosis of other organs} \]
\[ Y_{ij} = 1, \text{ if tuberculosis of other organs} \]

Then let us define the following model 1a:

\[ g(X_{ij}) = \beta_{0j} + \beta_{1j}\text{Infect} + \beta_{2j}\text{Nervous} \]

then the logistic regression model for (a) is

\[ P(Y_{ij} = 1 | X_{ij}) = \frac{e^{g(X_{ij})}}{1 + e^{g(X_{ij})}} \]

\[ P(Y_{ij} = 1 | X_{ij}) = \frac{e^{-5.690 + 5.133\text{Infect} + 2.749\text{Nervous}}}{1 + e^{-5.690 + 5.133\text{Infect} + 2.749\text{Nervous}}} \]

The results for logistic regression model are displayed in Table 4. It is evident from the table 4 that diseases of the nervous system (p \( \leq \) 0.25) and infection with parasite diseases (p \( \leq \) 0.25) have been positively associated with the tuberculosis of other organs. We include the diseases of the nervous system in the model because of the clinical important.

| TABLE VI. Results of Logistic Regression Analysis |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| B               | S.E.            | Wald            | df              | Sig             | Exp(B)          |
| NERVOUS(1)      | 2.749           | 2.397           | 1.316           | 1               | 15.634          |
| INFECT(1)       | 5.133           | 1.156           | 19.727          | 1               | 169.524         |
| Constant        | -5.690          | 1.030           | 30.505          | 1               | 0.003           |

Dependent Variable: Tuber_Organ

The finding show that the odd of developing tuberculosis of other organs among patient who’s having diseases of the nervous system is 15.6 times of those non-diseases of the nervous system, and the odd of developing tuberculosis of other organs among patient who’s having Infection with parasite diseases is 169 times of those non-infections with parasite diseases

IV CONCLUSIONS

This paper examines the factors that influencing infection of tuberculosis to other organs. Tuberculosis of other organs may depend on factors related to infection with parasite diseases and Diseases of the nervous system. These two factors can predict the development of tuberculosis of other organs. This paper provides only a preliminary overview of the problems associated with the relationship with tuberculosis of other organs. The main purpose of this paper is to demonstrate different technique that can be employed to explain such relationship. Due to data limitations, some of the important variables could not be used. In this research paper, three different methods have been used:

(i) Regression models
(ii) Risk estimate using Chi-Square

(iii) Path Analysis

All the models take account of cross-section data. The third model uses the path analysis model for explaining the direct and indirect effect upon tuberculosis of other organs. First model uses regression analysis; it appears that among the age of Tuberculosis patients the most dominating factors in explaining infection of TB. The path models reveal the findings more explicitly due of involving all the variables. With more detailed data, the advantage of the covariate dependent path model will be more precise and obvious.

Amir et al., (2011) point out that the TB also have seven dominating factor related to age, gender, ethnic, area CD4 count, heterosexual and IVDU. According to them, the most influencing factor that contributes to the HIV-Infected Tuberculosis (TB) Patients are CD4 and IVDU. This study found that parasite infection among married patients is 2.7 times compared to the non-married patients, $\chi^2 = 3.990$, $p = 0.046$, $OR = 2.773$.

In the presence of the tuberculosis to other organs model, infectious with parasite diseases and diseases of the nervous system show significant association. The important identifying associated with the development of tuberculosis to other organs model was to be able to prevent the complication from setting in. On the basis of the results, findings indicated that the intervention by reducing diseases of the genito-urinary system and increase patient’s knowledge on how the TB can spread will give a tremendous impact to prevalence of tuberculosis to other organs in the population. These direct and indirect findings may provide us with better understanding of the process of tuberculosis to other organs.

REFERENCES