Mortality in Hospitalized Human Immunodeficiency Virus (HIV) and Tuberculosis (TB) Infection Patients: Zero Inflated Negative Binomial Death Rate (ZINBDR) Models

Mohd Asrul A.A\textsuperscript{1} and Nyi Nyi Naing\textsuperscript{2}

\textsuperscript{1}Statistic Department, Universiti Tun Hussien Onn 86400, Parit Raja, Batu Pahat Johor, Malaysia afendi@uthm.edu.my

\textsuperscript{2}Unit of Biostatistic and Research, Universiti Sains Malaysia 16150, Health Campus, Kota Bharu Kelantan, Malaysia naing@usm kk.edu.my

Abstract

There has been considerable research conducted over the last 20 years focused on mortality in various cases of disease. The range statistical models commonly applied includes binomial, Poisson, negative binomial and zero inflated distribution. Negative binomial was used where Poisson distribution in analysis existing overdispersion. Such Poisson distribution, negative binomial is also widely used in medical studies, epidemiological, aquaculture and accident analysis and any related descriptive analysis. When there is an excess number of zero count useful approach to used a mixture model with a proportion $\omega$ of subjects not at risk, and a proportion of $1-\omega$ at risk subject who take on outcome values following negative binomial distribution. Explanatory variables effects can be incorporated into both components in the models. In HIV+TB assessment, mortality rate is highly recorded and full consideration is needed in this problem. The purpose of this study was to identify clinical and epidemiological factors associated with death in patients with an in-hospital diagnosis of human immunodeficiency virus (HIV) / acquired immune deficiency syndrome (AIDS) tuberculosis (TB), in a city with a high prevalence of TB and human immunodeficiency virus (HIV) infection. The risk of mortality once patients with TB are hospitalized
is unlikely to be explained only by the HIV epidemic. Death rate in HIV+TB assessment during period between 2001-2009 for Kelantan state, Malaysia.

**Keywords**: Zero Inflated Negative Binomial Death Rate, Standardization Rate, HIV/AIDS and TB

### 1 Introduction

To handle analysis involving count observation Poisson is a basic technique was used and overdispersion is widely seen in this particular. When a data consists zero values zero inflated poisson (ZIP) is appropriate model to replace. In such cases, when ZIP model existing overdispersion and highly assessing zero values, zero inflated negative binomial (ZINB) is alternative methods that will be used \[1, 2, 3, 4, 5\]. Hence, when the overdispersion is the results of a negative binomial distribution specifically there is an excess number of zero counts and a mixture assigning a mass of \( \omega \) to be extra zero and a mass of \( 1 - \omega \) to unimodal negative binomial distribution will give better leading to ZINB model. This frame also facilitate the exploration of the disease processes.

The issue of human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS) and tuberculosis (TB) is global concern. It is a major cause of illness and death worldwide especially the countries has a low and middle income. HIV and TB is a common infection for the estimated 5.5 million South African living with HIV/AIDS (in a national population of some 48 million). Hence, the co-infection rate of HIV is estimated at 73% in all TB cases. The estimates incidence of TB in South African is 692 per 100,000 people \[6\] a rate the World Health Organization (WHO) classifies a a serious epidemic. Subsequently, these problem was estimated that one-third of the world population is infected with *Mycobacterium tuberculosis*.

In Kelantan, Malaysia is the city with the highest \((HIV + TB^+)\) incidence with co-infection rate 53.7% in 2001 until 2009. The percentage presence \((HIV + TB^-)\) is 45.9% with 56 cases. The objective of this study was to identify the clinical and epidemiological factors associated with death patients HIV with TB and without TB incidence. We using ZINB models via substitute rate function (age death rate) into a count observation to relate with death rate in zero inflated negative binomial rate of death (ZINBDR).

The structures of the articles as follows. In section 2, we present a methods and materials whereby a short review ZINBDR. In section 3 the results of
Mortality in hospitalized human immunodeficiency virus patients. This article ends with discussion and conclusion on improving the reference distribution of the results.

2 Methods and Materials

2.1 Mortality rates of standardization

Rates estimation was used in various disciplines such as percentage, crude birth rate, and fertility rate. Term rate refers to the amount of change occurring in a quality with respect to time. Basically in practice, rates refer to the amount of change in a variable over a specified time interval divided by the length of that interval. Hence, population data are used to study and compare statistical analysis of disease incidence, prevalence, and mortality. It usually indicates per 1000 or 100,000 by considering the actual previous population and should always be examined when assessing mortality of the population.

Thus, via with discrete distribution model for discrete method whereby specific rates of the study population. Such population of interest are utilized to a randomly chosen standard population. Death rate was obtained by dividing the expected number of events by the total standard population. The requirement for calculation of rates are as follows [7];

1. Number of events such as death, incident cases for each group in the study.
2. Population for each group in the standard population.
3. Constant population 1000 or 100,000.

2.2 ZINB Dependent Death Rate (ZINBDR)

To incorporate into ZINB regression model we employ a death rate function to dependent variable. Rate dependent variable are estimated by requirement as follows. Let’s assume mortality rate cases in the $j^{th}$ observation for $j = 1, 2, \ldots, n$ a categorical observation age rate death estimation, whereby supposed to be negative binomial distributed with $d_j$ is the expected death of rate cases. Age death rate normally was calculated using standard population rate as follows [7];

$$d_j = \frac{q_j e_j}{p_j}$$ (1)

where:
$q_j =$ Number of death among persons of a given age group.
\( p_j \) = Population of person in given age group in a standard population \\
\( e_j \) = Constant population.

The equation ZINBDR as follows;

\[
(Y_{dj} = y_{dj}) = \begin{cases} 
\omega_i + (1 - \omega_i)(1 + \psi \theta_i)^{-\psi^{-1}} & \text{for } y_{dj} = 0 \\
(1 - \omega_i)f(y; \theta, \psi) & \text{for } y_{dj} > 0 
\end{cases}
\]

(2)

and the log-likelihood ZINBDR is;

\[
L_c(y_{dj}; \gamma, \beta, \psi) = \sum_{y_{dj}=0}^{n} \ln[\exp(z_i' \gamma) + (1 + \psi \exp(x_i')^{-\psi})] + \sum_{y_{dj}>0}^{n} \sum_{j=0}^{y_{dj}-1} \ln(j + \psi^{-1}) + \sum_{y_{dj}>0}^{n} \left\{ -\ln(y_{dj})! - (y_{dj} + \psi^{-1}) + \ln(1 + \psi \exp(x_i' \beta)) + y_{dj} \ln(\psi) + y_{dj} x_i' \beta \right\} \\
- \sum_{dj=1}^{n} \ln[1 + \exp(z_i' \gamma)]
\]

(3)

2.3 Study design and location

We conducted a retrospective, cohort study at medical record Hospital University Sains, Malaysia (HUSM) is a general, tertiary care university-affiliated hospital medical record. It consists 177 HIV and TB patients beginning 2001 until 2009. Kota Bharu is surrounded by a medium area in Malaysia that encompasses eight residential areas. The ethics committee at HUSM approved access to patients' records. Patients confidentiality has been maintained.

2.4 Patients and data collection

We used information contained in the medical records HUSM to identify the cases of HIV and TB. This center is a database of the Kota Bharu areas which store information concerning all notifiable infectious and contagious diseases. We included only those patients who began treatment for HIV+TB after hospitalization.
We evaluated for medical records of 55 HIV patients with TB and 122 HIV patients without TB and predictors of mortality were assessed. The primary outcomes of this study were in the hospital mortality after discharge. The duration of follow-up period was 10 year. The following data were collected using a secondary data (standardized questionnaire) such as: Demographic area (sex, ethnic, smoking status, years) and clinical characteristic (CD4 cell count, blood transfusion (BT), Intravenous drug (IVDU)). Data after discharge were obtained from the medical records (HUSM) database.

3 Result

From year 2001 to year 2009, 55 patients HIV with TB and 122 patients without TB were hospitalized in databased medical record. Demographic and clinical characteristic of study population are given in table 1. Forty-three (79.6%) patients had smoking and eleven patients (20.4%) no smoking for HIV with Tb patients and the number HIV patients without TB is 78 (63.9%) and 44 (36.0%) is no smoking. Thirty-three (64.8%) patients had CD4 count obtained before TB treatment (in the preceding 3 months). The odd ratios CD4 count of these patients was 2.56 (CI=1.32- 4.98) and only 19 (35.2%) no CD4 count cells. It is because the CD4 count can be reduced by acute infectious like TB, this test was not requested during hospitalization.

Table 2 shows the results ZINBDR regression model for year 2001 until 2009 in Kelantan, Malaysia areas. P-values is less than $\alpha$, then the null hypothesis can be rejected and the parameter estimates is significant at the level. Refer to mortality rate ($HIV + TB^+$) with regard to the log-odds of being an extra zero the term sex was not significant ($P$-values $= 0.72; \alpha = 0.05$) set our $\alpha$ level to 0.05. We do not reject the null hypothesis followed by ethnic ($P$-values $= 0.50; \alpha = 0.05$).

Similarly with CD4 count, heterosexual and blood transfusion was not an independent predictor of death. Hence, term of smoking is statistically significant ($P$-values $= 0.03; \alpha = 0.05$) the null hypothesis will be rejected and concludes that the regression coefficient for smoking term different from a zero given the other variables in the model. In NBDR part, all independent predictor are significant ($P$-values $= 0.00; \alpha = 0.05$). The data indicate no overdispersion once the extra zero were allowed.

Subsequently, for mortality patient $HIV + TB^-$ also shows the smoking term was significant ($P$-values $= 0.04; \alpha = 0.05$) and others independent predictors of death are not significant. We did not include residential areas in the analysis because it was not consider as a cause factor for the mortality cases.
Figure 1 and 2 shows the Kaplan Meier survival curves for 54 HIV patients with TB for smoking and no smoking cases during in-hospital length of stay.

Figure 1: Kaplan Meier survival curves for 11 HIV patients with TB for smoking cases during in-hospital length of stay.

4 Discussion

Zero inflated model have been used in medical research, traffic safety studies for modelling crashes for various application. Although zero inflation models offer improved statistical fit in many area of interest, it is argued in this paper that inherent assumption of a dual state process underlying the development of
Mortality in hospitalized human immunodeficiency virus

these model inconsistent with mortality or crash data. If ZIP and ZINB models provide good statistical fit but do not characterize the underlying mortality process, several interesting questions arise. First, what are the consequences of applying these models? Second what alternative model should be use instead? The answer for two this question, are if the only goals consists of finding the best statistical fit than the zero inflated model may be appropriate, since they improved statistical fit compared to NBDR or ZINBDR [3].

In this retrospective study, we demonstrated a high mortality rate especially in Kota Bharu areas 29 (53.7%) HIV with TB positive and 56 (49.5%) HIV with TB negative. Tb death are crucial indicator in TB programme monitoring [8], especially area with high HIV and TB prevalence such in Kelantan. The data on TB death provide us with a better understanding of the cause of these deaths and help guide interventions to reduce mortality. To our knowledge there are small studies that examined the risk factor for mortality in hospitalized patients HIV with and without TB in Kelantan. We using ZINBDR to found a strong association between age death rate and factor mortality death rate.

Several studies [9, 10, 11] found simiarly high mortality rates, particularly in patients TB-destroyed lungs and so on. Unlike other studies we did not find and association between a diagnosis of HIV and mortality. Thus the CD4 count (available for 64% of HIV + TB + patients) was not independent predictor in our study. The lack association between high mortality rate in hospitalize TB patients and HIV epidemic in our study could be explained by fact that like HIV population, or our non HIV population had a high proportion of comorbidities and were severely will high mortality rates.

These study has some limitations. Firstly the investigation was done in the single center. Secondly, retrospective study design did not allow us to assess some potential risk factor for mortality. Hence, these results suggest that a countinouse process of investment education and epidemiology surveillance of TB in hospital. Our data also suggest that being a smoker is an independent predictor of mortality after discharge. Both smoking and TB are still major causes of death, especially in developing countries. Smoking is a potentially modifiable risk factor. Therefore, prevention of death and TB control should include smoking related interventions.

As a conclusion, this study identified a high overall mortality rate among patients hospitalized with HIV and TB patients. In addition, the risk mortality once patients with HIV and TB are hospitalized is unlikely to be explained only by the patients with HIV-TB epidemic. The risk factor associated with death
after discharge were being a smoker and total lenght of hospital stay. Moreover, prospective studies are warranted to confirm these data and to identify possible interventions that can be reduce the mortality of patients admitted to hospital.

Acknowledgement. We would like to thank the Ministry of Higher Education and Universiti Tun Hussien Onn Malaysia, were supported this research. Thanks to medical record Hospital Universiti Sains Malaysia provided the data throught ethical approved on 15 Jun 2011 (USMKK/PPP/JEPeM [238.4(1.12)])

References


Received: September, 2012
Table 1: Demographic and clinical characteristic HIV with TB and without TB co-infection patients.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>HIV + TB</th>
<th>HIV + TB</th>
<th>RR (95% CI)</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>34±(6.86)</td>
<td>32.6±(11.01)</td>
<td>-</td>
<td>2.06(0.79-5.36)</td>
</tr>
<tr>
<td>Gender</td>
<td>48(88.9%)</td>
<td>97(79.5%)</td>
<td>0.89(0.79-1.02)</td>
<td>1.84(0.80-4.24)</td>
</tr>
<tr>
<td>Ethnic</td>
<td>7(13.0%)</td>
<td>111(91.0%)</td>
<td>1.05(0.93-1.18)</td>
<td>0.70(0.29-1.70)</td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4</td>
<td>35(64.8%)</td>
<td>51(41.8%)</td>
<td>0.65(0.48-0.86)</td>
<td>1.65(1.12-2.45)</td>
</tr>
<tr>
<td>Heterosexual</td>
<td>14(25.9%)</td>
<td>28(23.9%)</td>
<td>0.86(0.51-1.54)</td>
<td>1.90(0.89-4.04)</td>
</tr>
<tr>
<td>Smoking</td>
<td>43(79.6%)</td>
<td>78(63.9%)</td>
<td>0.80(0.66-0.97)</td>
<td>1.70(0.99-3.16)</td>
</tr>
<tr>
<td>Blood Transfusions</td>
<td>43(79.6%)</td>
<td>78(63.9%)</td>
<td>0.80(0.66-0.97)</td>
<td>1.70(0.99-3.16)</td>
</tr>
<tr>
<td>Residential areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Zero Inflated Negative Binomial Death Rate with α = 0.05
Figure 2: Kaplan Meier survival curves for 11 HIV patients with TB for no smoking cases during in-hospital length of stay.