

## **Original Research Article**

MODELLING THE INFLUENCE OF COVARIATES ON ADMISSION CASES OF TYPHOID PATIENTS AT THE TAMALE TEACHING HOSPITAL: AN APPLICATION OF GENERALISED LINEAR MODELLING CONCEPT

### ***Abstract***

Typhoid fever and typhoid perforation has become a major burden and causes a lot of deaths in the world especially in developing countries. This study considered a total of 418 patients diagnosed of Typhoid fever at the Tamale Teaching Hospital covering the period January 2010 to June 2015.

The data was modelled using the Negative binomial regression model to ascertain the duration (days) of admission cases as well as the effect of some covariates. Preliminary analysis revealed children of school going age as the most vulnerable. Further, collinearity diagnostics did not exhibit interaction among the predictors. The results identified cost of treatment, location of patients and outcome of admission as significant contributing covariates.

Based on the results, it was recommended that policy makers should take keen interest on the cost of treatment of typhoid since rising cost of treatment is likely to increase the number of days spent in the hospital thereby putting pressure on hospital facilities.

**Key words:** Negative Binomial, Poisson Regression, Salmonella, Typhoid and Covariates

### **1.0 Introduction**

Typhoid fever is one of the leading causes of morbidity and mortality across the world (Nagashetty et al., 2010). Typhoid is caused by a bacterium of the genus *Salmonella*. *Salmonella* infection in humans can be categorised into two broad types, that caused by low virulence serotypes of *Salmonella enterica* which cause food poisoning, and that caused by the high virulence serotypes *Salmonella enterica typhi* (*S. typhi*), that causes typhoid, and a group of serovars, known as *S Paratyphi* A, B and C, which cause Paratyphoid (Kanungo et al., 2008). Humans are the only host of this latter group of pathogens. *S. Typhi* is a highly adapted human-specific pathogen (Bhan et al., 2005), and the illness caused by these bacteria is a serious public health concern, particularly in developing countries (Naheed et al., 2010).

However, typhoid and paratyphoid fevers are still common in less-industrialized countries, owing to the consumption of unsafe drinking water, inadequate sewage disposal and flooding. The last

important outbreak recorded by the World Health Organization (WHO) counted a total of 42,564 cases and 214 deaths (case fatality ratio, 0.5%) in the Democratic Republic of the Congo from September 2004 to January 2005 (WHO Epidemic and Pandemic Alert and Response, 2008). The annual incidence of typhoid is estimated to be about 17 million cases worldwide, resulting in 500,000 to 600,000 deaths, mainly in endemic areas. However, the fact that countries across the world are more interdependent than ever implies that typhoid fever is still observed in every country, every year: in the US, about 400 cases occur each year (Centers for Disease Control and Prevention CDC, 2008). Typhoid fevers are systemic diseases characterized by diarrhea, fever, loss of appetite, headache, malaise, abdominal pain and progressive weakness, prostration, confusion and delirium. When untreated, typhoid fever persists for three weeks to a month; the most common treatment consists of antibiotics such as ampicillin, amoxicillin or ciprofloxacin. Possible complications that can lead to death are intestinal hemorrhage and perforation, kidney failure and peritonitis (Mathura et al., 2003).

Typhoid is transmitted by the ingestion of food or water contaminated with human feces from an infected person (Mathura et al., 2003). A few people can become carriers of *S. typhi* and continue to shed the bacteria in their feces for years, spreading the disease, as occurred in the famous case of “Typhoid Mary,” a cook who was closely associated with forty-seven cases and three deaths in New York over 100 years ago (Myron et al., 2007). Diagnostic suspicion is based on clinical data complemented by information on travel. For the final diagnosis to be established, the bacteria have to be detected in stool, blood or bone marrow samples, or by the Widal test. There are two vaccines currently recommended by the WHO, both available on the Spanish market. One is the oral Vivotif manufactured with live mutant bacteria of *Salmonella typhi* Ty21a, and the other is the injectable Typhim which consists of the Vi capsular polysaccharide (Germanier and Fuer, 1975 and Parry et al., 2002). Both are protective against *Salmonella typhi* infection and are recommended for travelers going to areas where typhoid is endemic. There are several studies that show crossprotection against *Salmonella paratyphi* A and B with Vivotif (Levine et al., 2007 and Tagliabue et al., 1986). In Spain, the primary risk population consists of travelers to the Indian

Subcontinent and Africa (Cook et al., 2007).<sup>10</sup> Typhoid and paratyphoid fever are important causes of morbidity and mortality and they result in large financial expenditures for the National Health Care System, especially due to hospitalizations (Todd, 1989).

A recent estimate found that 22 million new typhoid cases occur each year in the world with some 200,000 of these resulting in death (Crump et al., 2004), indicating that the global burden of this disease has increased steadily from a previous estimate of 16 million (WHO, 1997) however, case-fatality rates have decreased markedly (Crump et al., 2004). The highest number of cases (>100 per 100,000 persons/year) and consequent fatalities are believed to occur in South Central and Southeast Asia (Nagashetty et al., 2010). Generally, typhoid is endemic in impoverished areas of the world where the provision of safe drinking water and sanitation is inadequate and the quality of life is poor.

Although contaminated food (Sharma et al., 2009; Ram et al., 2007; Luby et al., 1998; Velema et al., 1997 and Black et al., 1985) and water (Luby et al., 1998), (Kelly-Hope et al., 2007; Gasem et al., 2001; Mermin et al., 1999 and King et al., 1989) have been identified as the major risk factors for typhoid prevalence, a range of other factors have been reported in different endemic settings such as poor sanitation (Karkey et al., 2010), close contact with typhoid cases or carriers (Tran et al., 2005), level of education, larger household size, closer location to water bodies (Tran et al., 2005 and Sur et al., 2007), flooding (Vollaard et al., 2004), personal hygiene (Kelly-Hope et al., 2007), poor life style (Kothari et al., 2008), and travelling to endemic areas (Whitaker et al., 2009). In addition, climatic variables such as, rainfall, vapour pressure and temperature have an important effect on the transmission and distribution of typhoid infections in human populations (Kelly-Hope et al., 2007 and Wang et al., 2012).

According to Marks *et al.*, (2010) typhoid fever ranks among the leading 20 causes of outpatient illness in Ghana, accounting for 0.92% of hospital admissions. The results raised an alarm since over 60% of the food vendors were infected. There was a higher risk of infecting those who patronise their food. This high figure according to them was due to lack of portable water, good

personal and environmental hygiene. Marks *et al.*, (2010) also posited that typhoid infection among children below five years was high and identified defecating in free range as the main contributing factor in the Agogo community of the Ashanti region of Ghana.

There has been very scanty scientific studies on the issue of typhoid fever in the northern Ghana. The issue of key covariates which turn to influence the length of stay of typhoid patients at the hospital is also been neglected in the scanty literature available in Ghana. ~~It is against this background that, this study attempts to model the covariates that influence the length of stay of patients with typhoid fever in the hospital.~~

## **2.0 Materials and Methods**

### **2.1 Data Management and Sample Size**

The data used in this study was a secondary data from Tamale Teaching Hospital. The sample size was four hundred and eighteen (418). It covers the period from January 2010 to June 2015. This data was extracted from among other ailments from the District Health Information Management System (DHIMS). Tamale has a total population of 2,468,557 people (Population Census, 2010). A design matrix was formed after the extraction of the data, and the variables captured included the following, gender (sex), outcome of admission, principal diagnosis, location of patient (distance), occupation of patient, insurance of patient, ward admitted, age of patient, cost of treatment, date of admission and date of discharged (duration).

These variables were all coded as, gender ( female = f, male = m), outcome of admission was categorized into three, either patient was discharged, absconded or died (discharged = d, absconded = a, death = t) , diagnosis of patients were into two categories (Typhoid Fever = t, Typhoid Perforation = p), location of patient was considered in terms of distance of the patient from his/her residence to the hospital. This was classified into two categories and coded as ( less than 10 kilometers = 1, greater or equal to 10 kilometers = 2), insurance status of the patient was also categorized into two and coded as yes = insured, no = not insured. Ward of admission was also coded as gynecology = g, medical = m, pediatrics = p, surgical = s, maternity = w. Age of patient was coded as under age (0-17years) =1, work force (18-60years) =2, aged (61years and

above) = 3. Cost of treatment and duration were continuous variables.

## 2.2 Concept of Generalized Linear models and Model building

Generalized Linear Model (GLM) is considered to be a flexible generalization of ordinary linear regression which allows for response variables that have error distribution models other than a normal distribution. The GLM generalizes linear regression by allowing the linear model to be connected to the response variable through an interaction and by allowing the magnitude of the variations of each measurement to be a function of its predicted value. Generalized linear models was formulated by John & Robert (1972) as a way of uniting various other statistical models, including linear regression, logistic regression and Poisson regression. They proposed an iteratively reweighted least squares method for maximum likelihood estimation of the model parameters. Maximum-likelihood estimation remains popular and is the default method on many statistical computing packages. Other approaches, including Bayesian approaches and least squares fits to variance stabilized responses, have been developed. There are several assumptions such as normality, homoscedasticity and linearity. However, in Generalized Linear models, normality assumptions is usually dropped for exponential distributions, homoscedasticity is also dropped for the variance function, and the linear assumption is also dropped for the link function and then translate the nonlinearity into a function of linear relationships which we call linear predictors.

A generalized linear models is made up of a linear predictor

$$Y = B_0 + B_1x_1 + B_2x_2 + \dots + B_px_p \quad [1]$$

and two functions

$$1. \text{ A link function describes how the mean, } E(Y_i) = u_i \quad [2]$$

depends on the linear predictor

$$g(u_i) = Y_i \quad [3]$$

$$2. \text{ A variance function that describes how the variance, } var(Y_i) \quad [4]$$

140 depends on the mean

$$141 \quad (var(Y_i) = \phi v(\mu) \quad [5]$$

142 Where the dispersion parameter  $\phi$  is a constant (Turner 2008)

143 The model of our study was given as

$$144 \quad \log Y = B_0 + B_1 x_1 + B_2 x_2 + B_3 x_3 + \dots + B_n x_n \quad [6]$$

145 The study looked at the Poisson distribution since a count data was considered. A negative  
146 binomial was subsequently fitted after the assumptions underlying the Poisson regression were not  
147 satisfied.

### 148 **2.2.1 Poisson Distribution**

149 In analysing the data using Poisson distributions, it assumptions must be considered. And when  
150 such assumptions are violated then there will be a need to cure those violations in order to have an  
151 appropriate model. If either of these assumptions is violated, they can lead to extra variation,  
152 sometimes referred to as over dispersion. These assumptions are: The probability of observing a  
153 single event over a small interval is approximately proportional to the size of that interval, the  
154 probability of two events occurring in the same narrow interval is negligible, the probability of an  
155 event within a certain interval does not change over different intervals and the probability of an  
156 event in one interval is independent of the probability of an event in any other non-overlapping  
157 interval.

158 The number of days each patient stay at the hospital (in days) is recorded as count. As a

159 consequence, the Poisson regression model is particularly appropriate for this type of response.

160 The model is non-linear and describes the mean number of days. The distribution of the number

161 of days using a Poisson model is of the form:  $\text{Prob}(Y=y) = \frac{e^{-\lambda T} (\lambda T)^y}{y!}$ ,  $y = 0, 1, 2, \dots$

$$162 \quad [7]$$

163 Where  $y$  represents the observed number of days,  $\lambda$  is a non-linear regression defines by

$$164 \quad \lambda = e^{\beta x} > 0 \quad [8]$$

165 The model matrix  $X$  is a  $(p \times 1)$  vector of explanatory variables and  $\beta$  is a  $(p \times 1)$  vector of  
 166 regression parameters. The conditional mean  $E(y/x) = \lambda = e^{\beta'x}$  [9]

167 The variance, (also called the heteroskedastic conditional variance) of the random variable is  
 168 constrained to be equal to the mean i.e  $\text{var}(y/x) = \lambda$  [10]

169 In most practical situations such as ours, the estimates in (8), namely the  $\beta$ 's are estimated by the  
 170 method of maximum likelihood, although different methods are also available (Hilbe and Greene  
 171 2008; Greene 2003; and Wooldridge 2003). Because of the equidispersion assumption which, in  
 172 most cases, is typical, our analysis added a fit of a more general specification model known as  
 173 the Negative Binomial (NB) model. The mean of the Poisson distribution is  $\lambda$ . For the Poisson  
 174 distribution, the variance,  $\lambda$ , is the same as the mean, so the standard deviation is  $\sqrt{\lambda}$ .

## 175 **2.2.2 Negative Binomial distribution**

176 This analysis is considered to be a fit of a more general specification model known as the  
 177 Negative Binomial (NB) model. This model (at least one alternative of) has become the standard  
 178 choice for basic statistical analysis of count data (Greene, 2008) especially among statisticians  
 179 i.e., the so-called generalized linear model. (Anscombe, 1949) is among the first to use this as a  
 180 tool for the analysis of count. A negative binomial experiment has the following properties: (a)  
 181 The experiment consists of  $x$  repeated trials (b) Each trial can result in just two possible outcomes.  
 182 We call one of these outcomes a success and the other, a failure. (c) The probability of success,  
 183 denoted by  $P$ , is the same on every trial.(d) The trials are independent from one another; the  
 184 outcome on one trial does not affect the outcome on other trials.

### 185 **2.2.2.1 Mean of the Negative Binomial Distribution**

186 If we define the mean of the negative binomial distribution as the average number of trials required  
 187 to produce  $r$  successes, then the mean is equal to:  $\mu = \frac{r}{P}$

188 Where  $\mu$  is the mean number of trials,  $r$  is the number of successes, and  $P$  is the probability of a  
 189 success on any given trial. The Negative Binomial probability distribution of  $Y$  is

$$Prob\left(\frac{Y}{y}\right) = \left(\frac{(r)}{(r + \lambda)}\right)^r * \frac{\Gamma(r + \lambda)}{\Gamma(y + 1)\Gamma(r)} * \left(\frac{\lambda}{\lambda + r}\right)^y, \quad y = 0, 1 \dots \quad [11]$$

Where  $\Gamma$  is the gamma function ( Kiebel and Holmes, 2003). In this case, the mean of the negative binomial distribution is the same as that of the Poisson i.e.

$$E\left(\frac{y}{x}\right) = \lambda \quad [12]$$

$$\text{But the variance is } Var\left(\frac{y}{x}\right) = \lambda + \frac{\lambda^2}{r} \quad [13]$$

Where  $r$  is the so-called dispersion parameter. The conditional variance of the Negative Binomial distribution exceeds the conditional mean. A general class of Negative Binomial is given in (Cameron and Trivedi (1986), and Greene (2008). The variance function is  $\mu_i + \alpha\mu_i p$

The relationship in equation 13 is the variance of the negative binomial model, where  $r$  is the dispersion parameter. A larger value of  $r$  reduces the NB to Poisson. This is an establishment that clearly shows the correlation between the two formular.

In classical linear model,  $\varepsilon$  is considered to obey the normal distribution with the involvement of the identity link between  $Y$  and  $\mu$ , the poisson regression uses the log link function to link the number of days spent in hospital. The negative Binomial uses the log-gamma link to establish the relationship in the negative binomial.

### 3.0 Results

Interaction among the independent variables need to be checked regarding generalized linear models, in particular the Poisson model that was employed here. This was done through collinearity diagnostic test using the tolerance and the variance inflated factor (VIF) values to assess the interaction. From Table 1, it is seen that, the independent variables did not exhibit any interaction. This is evident as the VIF figures for each independent variable is within the reference value of  $1 < VIF < 5$ . Also using the tolerance values from the Table 1 shows no interaction since

each value is less than the reference figure of 1.

**Table 1 Collinearity Diagnostic**

| Model Variables        | Collinearity Statistics |       |
|------------------------|-------------------------|-------|
|                        | Tolerance               | VIF   |
| Cost of Treatment      | 0.788                   | 1.268 |
| Gender of Patient      | 0.947                   | 1.056 |
| Outcome of Admission   | 0.942                   | 1.062 |
| Principal Diagnosis    | 0.922                   | 1.085 |
| Location of Patient    | 0.881                   | 1.135 |
| Occupation of Patients | 0.864                   | 1.158 |
| Insurance of Patient   | 0.971                   | 1.029 |
| Ward Admitted into     | 0.754                   | 1.327 |
| Age of Patient         | 0.942                   | 1.062 |

The Poisson regression was used first in modeling the length of stay of typhoid patients since it was a count data. Goodness of fit test revealed a lack of fit hence the Negative Binomial Regression provided a good fit (See Table 2). The log link function was used to link the independent variables age, gender, outcome, speciality, insurance, location, cost, diagnosis and occupation.

**Table 2 Goodness of Fit Test for Poisson Regression and Negative Binomial**

| Parameters                           | Poisson  |     |       | Negative Binomial |     |       |
|--------------------------------------|----------|-----|-------|-------------------|-----|-------|
|                                      | Value    | DF  | Sig.  | Value             | DF  | Sig.  |
| Deviance                             | 1911.1   | 401 | 4.766 | 154.961           | 401 | 0.386 |
| Pearson Chi-Square                   | 2119.275 | 401 | 5.285 | 171.051           | 401 | 0.427 |
| Log Likelihood                       | -1747.94 |     |       | -1338.87          |     |       |
| Akaike's Information Criterion (AIC) | 3529.884 |     |       | 2711.731          |     |       |
| Bayesian Information Criterion (BIC) | 3598.487 |     |       | 2780.334          |     |       |
| Consistence AIC                      | 3615.487 |     |       | 2797.334          |     |       |

The Goodness-of-fit refers to the extent to which a fitted model will conform to the data used in the model construction. Table 2 above shows the goodness of fit for both Poisson regression and the Negative Binomial. It can be observed that the model selection indicators for Negative Binomial are much smaller than the indicators for Poisson, therefore among the two models Negative Binomial was best for the data set under this circumstance. The Chi-square and the

deviance values for Poisson was high compared to the reference value, which is an indication for *over dispersion*. The chi-square and the deviance values were 0.386 and 0.427 respectively. These values clearly indicate that, the model is well fitted since they both have values less than one.

**Table 3 Test of Model Effects for Negative Binomial**

| Source     | Wald<br>Chi-Square | DF | P-value |
|------------|--------------------|----|---------|
| Intercept* | 15.436             | 1  | 0.000   |
| Gender     | 0.366              | 1  | 0.545   |
| Outcome*   | 20.578             | 2  | 0.000   |
| Diagnosis  | 0.251              | 1  | 0.616   |
| Location*  | 5.309              | 1  | 0.021   |
| Occupation | 2.808              | 4  | 0.590   |
| Insurance  | 0.903              | 1  | 0.370   |
| Speciality | 3.997              | 4  | 0.406   |
| Cost*      | 96.653             | 1  | 0.000   |
| Age        | 0.128              | 1  | 0.721   |

**\*Sig. (P- value < 0.05)**

**Table 4 Negative Binomial Parameter Estimates**

| Parameter    | B      | Wild Chi-Square | DF | P-value |
|--------------|--------|-----------------|----|---------|
| Intercept*   | 1.665  | 4.255           | 1  | 0.039   |
| Gender=f     | 0.069  | 0.366           | 1  | 0.545   |
| Gender = f   | 0.000  | -               | -  | -       |
| Outcome=a    | -0.111 | 0.089           | 1  | 0.765   |
| Outcome=d*   | -0.881 | 20.576          | 1  | 0.000   |
| Outcome=t    | 0.000  | -               | -  | -       |
| Diagnosis=p  | -0.078 | 0.251           | 1  | 0.616   |
| Diagnosis=t  | 0.000  | -               | -  | -       |
| Location=1*  | -0.264 | 5.309           | 1  | 0.021   |
| Location=2   | 0.000  | -               | -  | -       |
| Occupation=a | -0.230 | 0.983           | 1  | 0.322   |
| Occupation=h | -0.474 | 1.334           | 1  | 0.248   |
| Occupation=s | -0.235 | 1.37            | 1  | 0.242   |
| Occupation=t | -0.409 | 2.035           | 1  | 0.154   |
| Occupation=u | 0.000  | -               | -  | -       |

|                     |        |        |   |       |
|---------------------|--------|--------|---|-------|
| Insurance=n         | 0.158  | 0.803  | 1 | 0.370 |
| Insurance=y         | 0.000  | -      | - | -     |
| Speciality=g        | 0.198  | 0.047  | 1 | 0.829 |
| Speciality=m        | 0.196  | 0.061  | 1 | 0.905 |
| Speciality=p        | 0.190  | 0.060  | 1 | 0.807 |
| Speciality=s        | 0.464  | 0.358  | 1 | 0.550 |
| Speciality=w        | 0.000  | -      | - | -     |
| Cost*               | 0.001  | 96.653 | 1 | 0.000 |
| Age                 | -0.031 | 0.128  | 1 | 0.721 |
| (Scale)             | 1      | -      | - | -     |
| (Negative Binomial) | 1      | -      | - | -     |

**\*Sig ( $P$ -value < 0.05)**

Table 4 shows the parameter estimates. For outcome of admission, the log of counts of days spent at the hospital is expected to decrease by 0.881 for every discharged case at the hospital holding all other significant predictors in the model constant. For cost of treatment, the expected log of counts of patient's days will increase by 0.001 for each 1 Ghana cedi increase in the cost of treatment of Typhoid patients holding all other significant predictors in the model constant. The log of counts of patients stay at the hospital whose residential location was less than 10 km is expected to decrease by 0.264 for a unit distance increased by a km while holding all other significant predictors in the model constant. This means that, the distance of typhoid patients contributes to how long they stay in the hospital. If the distance is 10km or more, the likelihood of patient staying longer in the hospital is high and if the distance is less than 10km, the number of days a typhoid patient stay in the hospital is likely to reduce. The log of odds of patient days is expected to increase by 1.665 days if the distance (less than 10 km) covered by patients, cost of treatment and death are all zero in the model.

It is also noted that, holding other variables constant, The log of counts of patients stay at the hospital whose occupation are into agriculture, handicraft, students and trade will reduce the expected length of stay by 0.23, 0.47, 0.24 and 0.41 respectively. When a patient is admitted into gynecology Ward, the log of counts of patients stay at the hospital increases the expected length of stay by 0.20 medical ward, pediatrics ward and surgical ward increases the length of stay by 0.20, 0.19 and 0.46 respectively. For the age of patient, the expected log of counts of patient's

stays will decrease by 0.31. In general it was observed from Table 4 that the variables that had significant effects ( $p < 0.05$ ) were outcome of admission, location of patient and cost of treatment.

#### 4.0 Discussions

This is the first typhoid covariates study, to our knowledge, for this region. Even though this concept has been employed for quite some time, in Ghana, these concept has been applied mostly to diarrhoea, cholera and malaria (Haque et al., 2012; Loth et al., 2010; Ali et al., 2002a & 2002b; Ali et al., 2003; Emch and Ali, 2001 and Emch, 2000). Although these studies have shown great potential in identifying the covariates for this disease, no study has been found in the literature that applied the same analytical tool to the study of length of stay of typhoid patients.

In common with previous findings (Naheed et al., 2010 and Saha et al., 2001), we found that the under age group of 0–17 recorded the highest cases of typhoid. This confirms that that the rate of infection is highest in young children suggest that the more emphasis in terms of the current vaccination policy. Interestingly, we noticed that typhoid was evident in all age groups, and disproportionately affects both sex of the population.

In the Tamale metropolis of Ghana, the male population is more exposed for working and other purposes than females, which may explain the higher infection rates obtained for the males in the population (Morris et al., 1984). Other factors such as location (that is either rural or urban) and cost of treatment may also be attributed to the disproportionate number of cases in the population (Karkey et al., 2010). Further study is therefore needed to determine the underlying risk factors, to prevent certain gender and age groups being infected by typhoid fever.

The mean and median age of the cases was found to be 17 and 14 respectively, which contrasts earlier findings in Dhaka (Ram et al., 2007) however it is similar to the results of studies from Pakistan, Indonesia and India (Sur et al., 2007; Vollaard et al., 2004 and Ochiai et al., 2008). A distinct seasonal variation was found with almost half (45%) of the reported cases found to have occurred during the raining season. This is contrasting to the finding of a prospective community-based study (Naheed et al., 2010) but supports other results (Sur et al., 2007 and Saha

et al., 2001).

## 5.0 Conclusions

Typhoid fever is an acute illness associated with fever caused by *Salmonella typhi* bacteria. It can also be caused by *Salmonella paratyphi*. ~~This paper considered a data from the period January 2010 to June 2015 from the Tamale Teaching Hospital. The main aim of this paper was to determine the average length of stay of typhoid patients at the hospital. At the end of the study, it was revealed that, the expected length of stay of typhoid patients was  $(10.48 \pm 0.119)$ . It also revealed that, the average cost of treatment for typhoid patient was 553.48 Ghana cedis. Generalized linear model was used to analyse the data. Both Poisson and Negative Binomial Distributions were employed. It was noted after both analysis that, Negative Binomial was most appropriate for the data. The test of the model effects indicated the variables that were significant. It shows that, the outcome of admission, location of patients and the costs of treatment were statistically significant ( $PV < 0.05$ ) in contributing to the length of stay of typhoid patients. The estimates of the parameters also showed the contributions of each variable in the work. It was realised that, a unit change in cost of treatment will increase the length of stay by 0.1% holding other variable constant. It was noted that, for keeping other variables constant, the outcome of admission (discharged) will reduce the expected length of stay by 88.1%. It can also be seen that, holding other variables constant, the location of patient will reduce the expected length of stay by 26%.~~

## References

1. Ali M, Emch M, Donnay JP, Yunus M, Sack RB, (2002a) The spatial epidemiology of cholera in an endemic area of Bangladesh. Soc Sci Med 55(6): 1015–1024.
2. Ali M, Emch M, Donnay JP, Yunus M, Sack RB, (2002b) Identifying environmental risk factors for endemic cholera: a raster GIS approach. Health Place 8(3): 201–210.
3. Anscombe F. J (1949). The statistical analysis of insect counts based on the negative binomial distribution. Biometrics 165-173.
4. Bhan M, Bahl R, Bhatnagar S (2005) Typhoid and paratyphoid fever. Lancet 366: 749–762.

5. Black RE, Cisneros L, Levine MM, Banfi A, Lobos H, et al. (1985) Case-control study to identify risk factors for paediatric endemic typhoid fever in Santiago, Chile. *Bull World Health Organ* 635: 899–904
6. Cameron C., Trivedi P., (1998). *Regression Analysis of Count Data*. Cambridge University Press, New York.
7. Centers for Disease Control and Prevention CDC [consulted 3/10/2008] Available at [http://www.cdc.gov/ncidod/dbmd/diseaseinfo/TyphoidFever\\_g.html](http://www.cdc.gov/ncidod/dbmd/diseaseinfo/TyphoidFever_g.html).
8. Cook FJ, Day M, Wain J, Ward LR, Threlfall EJ (2007). Cases of typhoid fever imported into England, Scotland and Wales (2000–2003) *Trans R Soc Trop Med Hyg* 2007; 101:398-404.
9. Crump JA, Luby SP, Mintz ED (2004) The global burden of typhoid fever. *Bull World Health Organ* 82: 346–353
10. Emch M, (2000) Relationships between flood control, kala-azar, and diarrheal disease in Bangladesh. *Environ Plan A* 32: 1051–1063.
11. Emch M, Ali M, (2001) Spatial and temporal patterns of diarrheal disease in Matlab, Bangladesh. *Environ Plan A* 33: 339–350.
12. Gasem MH, Dolamans WM, Keuter MM, Djokomoeljanto RR (2001) Poor food hygiene and housing as risk factors for typhoid fever in Semarang, Indonesia. *Trop Med Int Health* 66: 484–490
13. Germanier R, Fuer E.(1975) Isolation and characterization of gal E mutant Ty21a of *Salmonella typhi*: a candidate strain for a live oral typhoid vaccine. *J Infect Dis* 1975; 141:553-8.
14. Greene, W (2008). Functional forms of the negative binomial model for count data. *Economic letters* 99, 585-590.
15. Greenwood BM. *The microepidemiology of malaria and its importance to malaria control. Trans R Soc Trop Med Hyg* 1989;83 (Suppl):25-9.
16. Haque U, Scott LM, Hashizume M, Fisher E, Haque R, et al., (2012) Modelling malaria treatment practices in Bangladesh using spatial statistics. *Malar J* 11: 63.
17. Hilbe J. M, and W. Greene (2008). Count Response Regression Models. *Handbook of Statistics*, Vol 27 210-252.
18. Hilbe, J. (1994), “Log Negative Binomial Regression Using the GENMOD Procedure,” in *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute Inc.
19. Kanungo S, Dutta S, Sur D (2008) Epidemiology of typhoid and paratyphoid fever in India. *J Infect Dev Ctries* 26: 454–460

20. Karkey A, Arjyal A, Anders KL, Boni MF, Dongol S, et al. (2010) The burden and characteristics of enteric fever at a healthcare facility in a densely populated area of Kathmandu. *PloS One* 511: e13988.
21. Kelly-Hope LA, Alonso WJ, Theim VD, Anh DD, Canh DG, et al. (2007) Geographical distribution and risk factors associated with enteric diseases in Vietnam. *Am J Trop Med Hyg* 764: 706–712
22. Kiebel S.J and A.P. Holmes (2003). The General Linear Model. The General Linear model. In R.S.J. Frackowiak, K.J. Friston, C. Frith, R. Dolan, K.J. Friston, C.J. Price, S. Zeki, J. Ashburner, and W.D. Penny, editors, *Human Brain Function*. Academic Press, 2nd edition.
23. King CC, Chen CJ, You SL, Chuang YC, Huang HH, et al. (1989) Community-wide epidemiological investigation of a typhoid outbreak in a rural township in Taiwan, Republic of China. *Int J Epidemiol* 181: 254–260
24. Kothari A, Pruthi A, Chugh TD (2008) The burden of enteric fever. *J Infect Dev Ctries* 24: 253–259
25. Levine MM, Ferrecio C, Black RE, Lagos R, San Martin O, Blackwelder WC (2007).Ty21a live oral typhoid vaccine and prevention of paratyphoid fever caused by *Salmonella enterica* Serovar Paratyphi B. *Clin Infect Dis* 2007; 45:24-8.
26. Loth L, Gilbert M, Osmani MG, Kalam AM, Xiao X, (2010) Risk factors and clusters of highly pathogenic avian influenza H5N1 outbreaks in Bangladesh.
27. Luby SP, Faizan MK, Fisher-Hoch SP, Syed A, Mintz ED, et al. (1998) Risk factors for typhoid fever in an epidemic setting, Karachi, Pakistan. *Epidemiol Infect* 120(2): 129–138
28. Mathura KC, Gurubacharya DL, Shrestha A, Pant S, Basnet P, Karki DB (2003). Clinical profile of typhoid patients. *Kathmandu Univ Med J* 2003; 1:135-7.
29. Mermin JH, Villar R, Carpenter J, Roberts L, Samaridden A, et al. (1999). A massive epidemic of multidrug-resistant typhoid fever in Tajikistan associated with consumption of municipal water. *J Infect Dis*: 1416–1422
30. Morris JG Jr, Ferreccio C, Garcia J, Lobos H, Black RE, et al., (1984) Typhoid fever in Santiago, Chile: a study of household contacts of pediatric patients. *Am J Trop Med Hyg* 33(6): 1198–1202
31. Myron ML, Plotkin SA, Orenstein WA, Picazo JJ (2007).*Vacunas frente a la fiebre tifoidea*. editors. *Vacunas*. 4Th ed. Madrid: Editorial Médica AWE 2007; 1099-136.
32. Nagashetty K, Channappa ST, Gaddad SM (2010) Antimicrobial susceptibility of *Salmonella Typhi* in India. *J Infect Dev Ctries* 42: 070–073

33. Naheed A, Ram PK, Brooks WA, Hossain MA, Parsons MB, et al. (2010) Burden of typhoid and paratyphoid fever in a densely populated urban community, Dhaka, Bangladesh. *Int J Infect Dis* 14S3: e93–e99
34. Ochiai RL, Acosta CJ, Danovaro-Holliday MC, Baiqing D, Bhattacharya SK, et al., (2008) A study of typhoid fever in five Asian countries: disease burden and implications for controls. *Bull World Health Organ* 86(4): 260–268.
35. Parry CM, Hien TT, Dougan G, White N, Farrar J. Typhoid Fever *N Eng J Med* 2002; 347:1770-82.
36. *Prev Vet Med* 96: 104–113.
37. Ram PK, Naheed A, Brooks WA, Hossain MA, Mintz ED, et al. (2007) Risk factors for typhoid fever in a slum in Dhaka, Bangladesh. *Epidemiol Infect* 135: 458–465
38. Saha SK, Baqui AH, Abdullah H, Hanif M, Darmstadt GL, et al., (2001) Typhoid fever in Bangladesh: implications for vaccination policy. *Ped Infect Dis J*, 20(5), 521–524
39. Sharma PK, Ramakrishnan R, Hutin Y, Manickam P, Gupte MD (2009) Risk factors for typhoid in Darjeeling, West Bengal, India: evidence for practical action. *Trop Med Int Health* 146: 696–702
40. Sur D, Ali M, Seidlein LV, Manna B, Deen JL, et al. (2007) Comparisons of predictors for typhoid and paratyphoid fever in Kolkata, India. *BMC Public Health* 7: 289.
41. Tagliabue A, Villa L, De Magistris MT, Romano M, Silvestris S, Boraschi D, et al. (1986). IgA-driven T cell-mediated anti-bacterial immunity in man after live oral Ty21a vaccine. *J Immunol* 1986; 137:1504-10.
42. Todd EC (1989). Costs of acute bacterial foodborne disease in Canada and the United States. *Int J Food Microbiol* 1989; 9:313-26.
43. Tran HH, Bjune G, Nguyen BM, Rottingen JA, Grais RF, Guerin PJ (2005) Risk factors associated with typhoid fever in Son La Province, northern Vietnam. *Trans R Soc Trop Med Hyg* 99: 819–826
44. Velema JP, van Wijnen G, Bult P, van Naerssen T, Jota S (1997) Typhoid fever in Ujung Pandang, Indonesia- high risk groups and high-risk behaviours. *Trop Med Int Health* 211: 1088–1094
45. Vollaard AM, Ali S, van Asten HAGH, Widjaja S, Visser LG, et al. (2004) Risk factors for typhoid and paratyphoid fever in Jakarta. *JAMA* 29121: 2607–2615
46. Wang LX, Li XJ, Fang LQ, Wang DC, Cao WC, et al. (2012) Association between the incidence of typhoid and paratyphoid fever and meteorological variables in Guizhou, China. *Chinese Med J* 1253: 455–460]

- 412 47. Whitaker JA, Franco-Pardes C, del Rio C, Edupuganti C (2009) Rethinking typhoid fever  
413 vaccines: implications for travelers and people living in highly endemic areas. J Travel Med  
414 161: 46–52
- 415 48. WHO Epidemic and Pandemic Alert and Response [consulted 3/10/2008] Available at  
416 [http://www.who.int/csr/don/2005\\_01\\_19/en/index.html](http://www.who.int/csr/don/2005_01_19/en/index.html).
- 417 49. World Health organization (WHO), (1997). Typhoid fever. WHO, Geneva, Fact Sheet No.  
418 149.