

Modelling the Force of Infection for Hepatitis B among Heterogeneous Groups at the Komfo-Anokye Teaching Hospital

Frempong NK*, Nartey E, Bukari FK, and Berchie RO

Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

*Corresponding author: Frempong NK, Department of Mathematics, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, Tel: +233202999488, E-mail: nkf.cos@gmail.com

Citation: Frempong NK, Nartey E, Bukari FK, Berchie RO (2018) Modelling the force of infection for Hepatitis B among heterogeneous groups at the Komfo-Anokye Teaching Hospital. *J Biostat Biometric App* 3(1): 109

Abstract

The aim of this study is to estimate the age-dependent force of infection for hepatitis B virus (HBV) for sick patients showing symptoms, blood donors and people coming to the serology department to test for the antigen in blood voluntarily. A secondary data was obtained from the serology department at the Komfo Anokye Teaching Hospital (KATH, January 2013- June 2016). Linear and quadratic age-dependent beta regression models were fitted to the prevalence of HBV. From the results, the estimated quadratic model was suitable for the mean prevalence of HBV, with a significant positive linear and negative quadratic age effects respectively. From the results, the prevalence increases from lower ages to age 38 and decreases at higher ages. The estimated force of infection for the heterogeneous group is high at lower ages, decreases to age 38 and increases to age 64. In conclusion, the prevalence is high at age 38 and the transmission rate is lower at age 38 but higher at age 11 and 64.

Keywords: Prevalence; Force of Infection; Beta Regression Model; Hepatitis B Virus

Introduction

Hepatitis B is a viral infection that attacks the liver and is caused by the hepatitis B virus (HBV). Globally, in 2015, an estimated 257 million people were living with chronic HBV infection. The epidemic caused by HBV affects mostly the WHO African Region and the Western Pacific Region (Global report 2017, W.H.O). In Africa with many developing countries, the highest rates of HBV carriers are due to primitive or limited medical resources. In these countries widespread of infection may occur in infancy and childhood. There are several sources of contracting the HBV namely Infants born to infected mothers, Health care workers, Sexually active heterosexuals, Injection drug users sharing unsterile needles, blood donors amongst others. Some studies have investigated the prevalence among these high risk groups of contracting the virus. HBV infection is a serious public health disease in sub-Saharan Africa with prevalence ranging between 3% and 22% in blood donors [1]. The HBV is said to be acute if it is short lived and may possibly have a cure. Chronic HBV infections are the persistent infection, and the likelihood of an acute infection becoming chronic is dependent on age. If one is infected by the virus in infancy, the probability of it becoming chronic is 0.90, [2].

Organization of this paper is as follows: Section 1 discusses the background of the study; Section 2 is where the validity of data and methodologies are discussed. Section 3 and 4 discusses results and conclusion respectively.

The prevalence of hepatitis B virus E antigen (HBeAg) among Ghanaian blood donors in the current study area (Komfo Anokye Teaching Hospital, KATH) have been studied by [3]. They reported an overall prevalence of 13.3%. Their results also revealed that the prevalence of HBeAg was higher in patients between the age group of 10-20 years and appeared to decrease with increase in age. A systematic review and meta-analysis of prevalence of HBV over a decade (1995-2015) revealed that across Ghana, highest HBV infection prevalence rates were recorded in persons within the age group 16-39 years [4]. From these studies, there is empirical evidence that HBV infections are dependent on age. The force of infection (FOI) which helps to measure the transmission of the HBV infection within a population has never being studied for this disease in Ghana. FOI is the rate at which exposed individuals are infected with HBV. Knowing the proportion of the population infected at a given time and the rate of transmission of HBV, improves the understanding of the HBV epidemic, and necessary for effective public health efforts. The main objective of this study is to model the prevalence of HBV infection as a function of age using Beta Regression technique and estimate FOI from the estimated model. The developed model can be used as a baseline for further research of the various risk groups [5].

Organization of this paper is as follows: Section 1 discusses the background of the study; Section 2 is where the validity of data and methodologies are discussed. Section 3 and 4 discusses results and conclusion respectively.

Materials and Methods

The data used was obtained from KATH from the period of January 2013 to June 2016 at the virology/ serology unit under the blood transfusion department. Overall, 18528 patients were involved. No duplicated records were reported over the period [6].

Beta Regression Model

The beta regression model is to model the mean response given age. The dependent variate in this study is the proportion of HBV+ at each age after infection [7].

Assumptions

The response is constraints to the standard unit interval (0,1) and the ages of the individuals is the age diagnosed just after infection. The proportion of HBV+ is beta distributed and lies in the flexibility of the assumed beta law. The model is naturally heteroskedastic and easily accommodate asymmetries. The beta density function is given by;

$$f(y : p, q) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} y^{p-1} (1-y)^{q-1}, \quad 0 < y < 1 \quad (2.1)$$

for $p, q > 0$ and $\Gamma(\cdot)$ is the gamma function. Here p represents the number of HBV+, q represent the number of HBV- and y is the observed proportion of HBV+. The mode exists when p and q are greater than 1.

The mean and variance of the proportion of HBV+ is given as follows:

$$E(Y) = \frac{p}{p+q}$$

$$Var(Y) = \frac{pq}{(p+q)^2(p+q+1)}$$

Let y_1, \dots, y_n be independent random variable where each $y_i, i=1, \dots, n$ follows the density with μ_i and unknown precision ϕ .

$$f(y : \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi} (1-y)^{(1-\mu)\phi-1} \quad (2.2)$$

$$0 < y < 1 \text{ with } 0 < \mu < 1 \text{ and } \phi > 0$$

$$\mu = \frac{p}{p+q} \quad \text{and} \quad \phi = p+q$$

Let y_1, \dots, y_n be a random sample such that $y_i \sim \beta(\mu_i, \phi), i = 1, \dots, n$.

The beta regression model is defined as:

$$g(\mu_i) = X_i^T \beta = \eta_i, \text{ where } \beta = (\beta_1, \dots, \beta_k)^T \quad (2.3)$$

is a $k \times 1$ vector of an unknown regression parameters, $k < n, x_i = (x_{i1}, \dots, x_{ik})^T$ is vector of k regressors and η is a linear parameter (ie., $\eta_i = \beta_1 x_{i1} + \dots + \beta_k x_{ik}$). From (2.3) usually $x_{i1} = 1$ for all i so that the model has an intercept. $g(\cdot) : (0,1) \rightarrow \mathfrak{R}$. We write $Y \sim \beta(\mu, \phi)$ where,

$$E(Y) = \mu \quad \text{and} \quad Var(Y) = \frac{\mu(1-\mu)}{1+\phi}$$

The parameter ϕ is known as the precision parameters since for fixed μ , the larger the ϕ the smaller the variance of Y ; ϕ^{-1} is a dispersion parameter.

Estimation

To estimate the parameter β , the log likelihood function based on a sample of n independent observations is

$$\ell(\beta, \phi) = \sum_{i=1}^n \ell_i(\mu_i, \phi)$$

where;

$$\begin{aligned} \ell_i(\mu_i, \phi) = & \log \Gamma(\phi) - \log \Gamma(\mu_i \phi) - \log \Gamma((1 - \mu_i)\phi) + (\mu_i - 1) \log(y_i) \\ & + (1 - \mu_i)\phi - \log(1 - y_i) \end{aligned} \quad (2.4)$$

With μ_i defined so that the model holds. The score function obtained by differentiating the log likelihood in (2.4) with respect to the unknown parameters is given by:

$$(u_\beta(\beta, \phi), u_\phi(\beta, \phi))^T \quad \text{where} \quad u_\beta(\beta, \phi) = \phi X^T T (y^* - u^*) \quad \text{with } X \text{ being } n \times k \text{ matrix whose } i_{th} \text{ row is } x_i^T, \\ T = \text{diag}(1/g'(\mu_1), \dots, 1/g'(\mu_n)), \quad y^* = (y_1^*, \dots, y_n^*)^T \text{ and } u^* = (u_1^*, \dots, u_n^*)^T$$

$$\text{var}(y_i) = \mu_i(1 - \mu_i)/(1 - \phi)$$

The maximum likelihood estimators of β and ϕ are obtained from the equation (2.5) and $\mu_\phi(\beta, \phi) = 0$ and do not have closed form. A global measure of explained variation is obtained by computing the pseudo $R^2(R_p^2)$ defined as the square of the sample correlations coefficient between η and $g(y)$. Note that $0 < R_p^2 < 1$ and a perfect agreement between η and $g(y)$ hence between μ and $g(y)$. It is also possible to define the standardized residuals;

$$r_i = \frac{y_i - \mu_i}{\sqrt{\text{var}(y_i)}}$$

Where; $\mu_i = g^{-1}(x_i^T \beta)$ and $\text{var}(y_i) = \mu_i(1 - \mu_i)/(1 - \phi)$

The model estimates the age specific force of infection of the prevalence of HBV by recognizing that the probability of infection at age x is the joint probability of infection at age x and the probability of not being infected in an interval 0 to $x-1$. The FOI at age x is defined as:

$$f_x = \frac{\pi'(x)}{1 - \pi(x)} \quad (2.5)$$

The logit function as a link function for the Beta model in (2.3) is defined as:

$$\pi(x) = \frac{e^{\beta^T x}}{1 + e^{\beta^T x}}$$

The model for quadratic age is shown here:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x + \beta_2 x^2}}{1 + e^{\beta_0 + \beta_1 x + \beta_2 x^2}} \quad (2.6)$$

differentiating (2.6) and substitute in (2.5) simplifies to:

$$\pi'(x) = \frac{(\beta_1 + 2\beta_2x)e^{\beta_0 + \beta_1x + \beta_2x^2}}{1 + e^{\beta_0 + \beta_1x + \beta_2x^2}}$$

$$1 - \pi(x) = \frac{1}{1 + e^{\beta_0 + \beta_1x + \beta_2x^2}}$$

Hence the force of infection is given as;

$$f_0 = \frac{(\beta_1 + 2\beta_2x)e^{\beta_0 + \beta_1x + \beta_2x^2}}{1 + e^{\beta_0 + \beta_1x + \beta_2x^2}} \tag{2.7}$$

Results and Discussion

A total of 18528 screened from year 2013 to 2016 registered at the virology/serology unit under the blood transfusion department were obtained. Of the 18528 screened 2574 (13.89%) were HBV+ and 15954 (86.11%) were HBV-. The ages of the subjects were categorized according to the 2010 Population and Housing Census (PHC) from the Ghana Statistical Service (GSS), where (12-19) represents the adolescent, (20-24) represents the young adults, (25-59) represents the old adults and lastly (60+) is the aged. The prevalence distribution for HBV infection stratified by 4 (four) age classes is shown in Table 1. HBV prevalence showed a decrease from 0.8% in 12-19 year olds to 11% in 25-59 year olds. It is observed that the prevalence of HBV is relatively higher among age group 25-59 and lower for the above 60 ages.

AGE GROUPS	HBV+	HBV-	TOTAL
12-19	161(0.8%)	943(5%)	1104
20-24	220 (1%)	1440(8%)	1660
25-59	2029(11%)	11537(62%)	13566
Above 60	164(0.8%)	2034(11%)	2198
TOTAL	2574	15954	18528

Table 1: Age groups of diagnosed patients

Figure 1 shows the proportion of positive samples per age class based on the serological data, with the size of the dots proportional to the number of samples collected in the corresponding age category. It can be observed that most infections are between age classes 22 to 45 years.

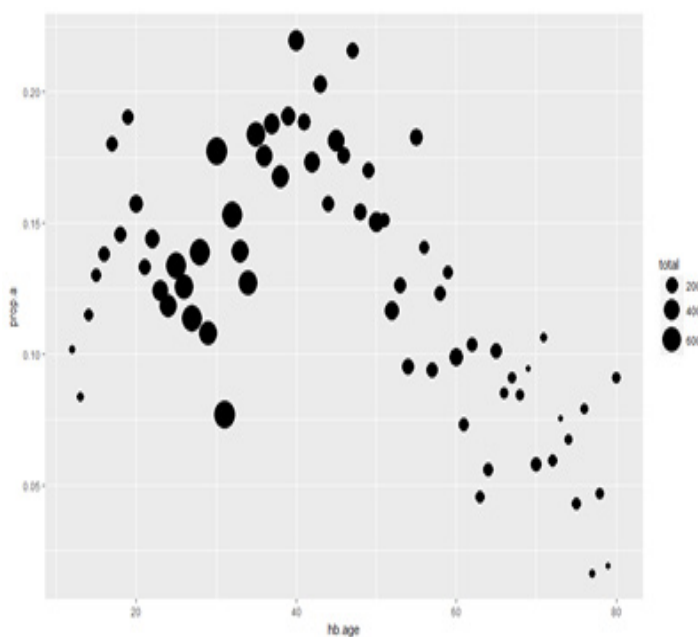


Figure 1: Age-specific prevalence of HBV with dots proportional to sample size

Figure 2 shows the prevalence of HBV at different single year age class. We observed that prevalence increases from adolescents and reach its peak at about age 40 and decreases for higher ages. It is observed that there is a steady growth and a sharp decline of HBV infections over ages as shown by the loess curve.

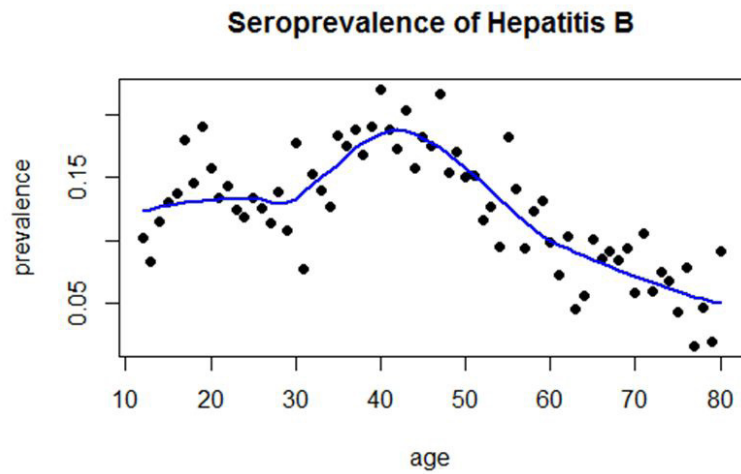


Figure 2: The loess curve on the prevalence of HBV

Figure 3 depicts a linear and quadratic age-dependent model for the prevalence using single year age groups. From the quadratic model, the prevalence of HBV depicts a concave shape. The linear model shows a negative slope for the prevalence with age. Results of analysis using both linear and quadratic age effects in the beta regression model is presented in Tables 2 and 3 respectively.

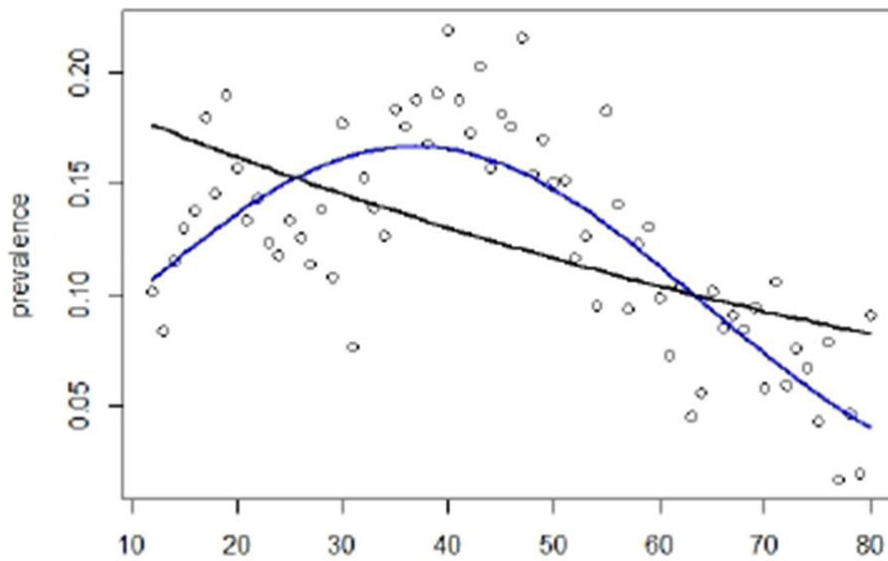


Figure 3: Linear and Quadratic model for Prevalence of HBV

In each case, results are presented for the fitted linear model and the quadratic model showing the parameter estimates, standard errors and significant test:

Coefficientss	Estimates	Standard error	z-value	p-value
<i>intercept</i>	-1.3851	0.1122	-12.34	<0.001
<i>hb.age</i>	-0.01282	0.0024	-5.34	<0.001
<i>Phi</i>	56.564	9.636		

Table 2: Parameter estimates of the linear model

Estimated linear Model:

$$g(\mu) = -1.385084 - 0.012818(hb.age)$$

Coefficients	Estimates	Standard error	z-value	p-value
<i>intercept</i> (β ₀)	-2.746	0.191	-14.361	<0.001
<i>hb.age</i> (β ₁)	0.062	0.0095	6.547	<0.001
<i>hb.sage</i> (β ₂)	-0.00084	0.00011	-7.990	<0.001
<i>phi</i> (φ)	114.54	19.51		

Table 3: Parameter estimates of the quadratic model

Estimated quadratic model:

$$g(\mu) = -2.746 + 0.062(hb.age) - 0.00084(hb.sage)$$

Results from both Tables 2-3 shows significant effect of age on the mean prevalence of HBV. The Akaike’s information criterion (AIC) for comparison of the two models was considered. The AIC of the linear model is -238.238 and that of the quadratic is -283.8155. The quadratic model had the least AIC hence was chosen for further analysis.

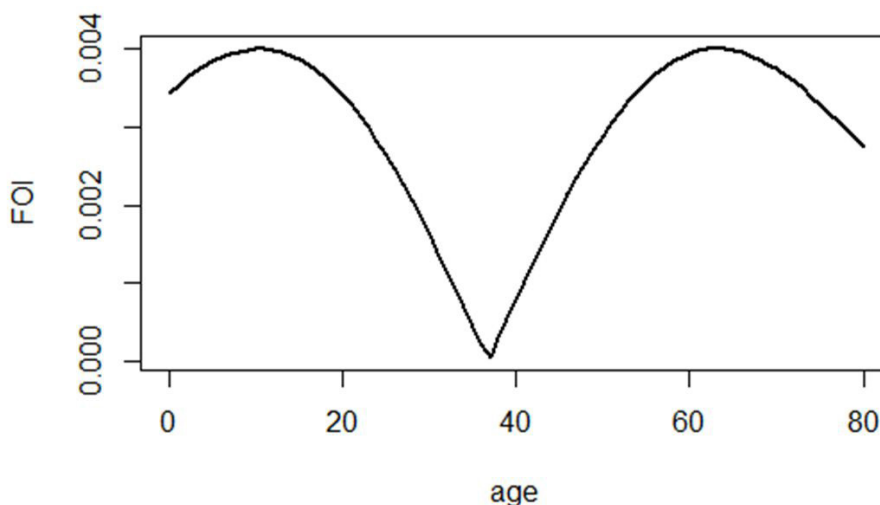


Figure 4: The estimated Force of Infection

The estimated quadratic model was used to estimate the FOI in equation (2.7). Figure 4 shows the estimated force of infection increases exponentially from age 0 to 12 and decreases from age 12 to age 38. The FOI increases from age 38 to 64 and decreases thereafter. An average force of infection (instantaneous per susceptible incidence) for HBV of 0.034%/annum, (i.e. 3.4/1000 at births/annum). However, age-dependence in transmission is supported statistically indicating peak incidence rate in the total population of 0.04%/annum in 11 year olds and 64 year olds, and a lower incidence of 0.005%/annum in 38 year olds.

Conclusion

HBV prevalence in Ghana has been studied by different authors. KATH is a hospital situated in a city in Ghana, so this result is not representative of Ghana in general. From the results, HBV infections increases from lower ages and reach its peak at age 38 years old. The historical HBV infections decrease at a faster rate from age 38 years to age 80 years old. In conclusion, the HBV prevalence is high at age 38 and the transmission rate is lower at age 38 years and higher at age 11 year and 64 year olds.

References

- Allain JP, Candotti D, Soldan K, Sarkodie F, Phelps B, et al. (2003) The risk of hepatitis b virus infection by transfusion in kumasi, ghana. *Blood* 101: 2419-25.
- Shepard CW, Simard EP, Finelli L, Fiore AE, Bell BP (2006) Hepatitis b virus infection: epidemiology and vaccination. *Epidemiol Rev* 28: 112-25.
- Rufai T, Mutocheluh M, Kwarteng K, Dogbe E (2014) The prevalence of hepatitis B virus E antigen among Ghanaian blood donors. *Pan Afr Med J* 17: No 1.
- Ofori-Asenso R, Agyeman AA (2016) Hepatitis B in Ghana: a systematic review & meta-analysis of prevalence studies (1995-2015). *BMC Infect Dis* 16: 130.
- Ferrari S, Cribari-Neto F (2004) Beta regression for modelling rates and proportions. *J Appl Stat* 31: 799-815.
- Ganem D, Prince AM (2004) Hepatitis B virus infection-natural history and clinical consequences. *New Eng J Med* 350:1118-29.
- Lavanchy D (2004) Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures. *J Viral Hepat* 11: 97-107.

Submit your next manuscript to Annex Publishers and benefit from:

- ▶ Easy online submission process
- ▶ Rapid peer review process
- ▶ Online article availability soon after acceptance for Publication
- ▶ Open access: articles available free online
- ▶ More accessibility of the articles to the readers/researchers within the field
- ▶ Better discount on subsequent article submission

Submit your manuscript at

<http://www.annepublishers.com/paper-submission.php>