Neural, Parallel, and Scientific Computations 25 (2017) 509-522

MODELING MATERNALINFANT HIV TRANSMISSION WITH LAG TIME DISTRIBUTIONS EXPONENTIAL, GEOMETRIC AND SHIFTED GEOMETRIC

SUBHASH SHENDE[1], MOHAN KALE[2], AND NIKHIL GUPTE[3]

[1]Department of Statistics, Fergusson College, Pune - 411004, India
[2]Department of Statistics, Savitribai Phule Pune University, Pune - 411007, India
[3]B.J. Medical College and Sassoon General Hospital, Pune - 411001, India

ABSTRACT. The main current public-health research question is whether breastfeeding by Human Immunodeficiency Virus(HIV)-infected mothers can be made safer as to transmission risk, given the possible adverse effects of refraining from breastfeeding. Various ongoing or planned trials and studies concentrates on mode of infant feeding or antiretroviral therapy to the infant over the breastfeeding period. An important task is to determine risk of transmission of perinatal HIV and when it occurs. Several ongoing HIV prevention trials throughout the developing world are evaluating different methods to reduce perinatal HIV transmission. Perinatal transmission refers to mother to infant HIV transmission occurring before or at the time of the birth. It results from fetal exposure to the maternal fluids or infected maternal secrations. The present article proposes statistical models that simultaneously estimates the risks of perinatal transmission together with the sensitivity of the screening tests for HIV infection. These models also allow estimating infectivity through breast feeding during postpartum period. The article aims at brief overview of various lag time distributions and presenting a tour of tools and techniques available in statistical literature for analysing such data sets. The proposed methodology is demonstrated with a case study.

KEY WORDS: HIV transmission probabilities; Mother-to-infant HIV transmission; Perinatal HIV transmission.

1. INTRODUCTION

An important public health issue is to determine the risk of transmission of perinatal HIV and when it occurs. Perinatal HIV transmission can occur either at antepartum, this is referring to the period from conception to delivery; at intrapartum, this is referring to the period during delivery; and postpartum, this is referring to the period following the birth. Perinatal transmission refers to, mother to infant transmission of HIV occurring before or at the time of the birth and results from fetal

Received January 31, 2017

exposure to the maternal fluids or infected maternal secretions. Several hypothesis regarding the mechanism and timing of transmission have been proposed, although the exact mechanism remains unknown. It is estimated that 90 percent HIV infected infants are infected through perinatal transmission (Raji Balasubramanian et al.,2001). Several ongoing HIV prevention trials through out the developing world are using different methods to reduce perinatal HIV transmission. According to United Nations Programme on HIV/ AIDS, an estimated 1800 HIV infected infants are born each day in the developing world. Various studies have shown that treatment can reduce the transmission rate by 67 percent or more (Conor, E.M.et al., 1994). In 2010, 1.8 million people died from HIV/AIDS and another 2.6 million people were reported to be infected with the virus. The risk of HIV infection during the postpartum period arises when an HIV infected mother breastfed her infant. Estimation of the distribution of the time of perinatal transmission is difficult because tests of infection status can only be undertaken after birth. DNA and RNA polymerase chain reaction (PCR) assays and HIV culture have been most commonly used as diagnostic tests for perinatal HIV infection. The risk of transmission by an infected mother occurring before or during birth (without interventions to reduce the transmission) is 15-25 percent (A review of available evidence 2004). The present study is an extension of N.Gupte et al., 2007 work carried out as a randomized control study, conducted in Johannesburg, South Africa. This was an open label clinical trial of the use of Nevirapine (NVP) and Zidvudine (AZT) administered postnatally to infants born to HIV infected mother who had no prior treatment with antiretroviral theraphy (ART).

The article is organised as follows. Section 2, is an introductory material giving the data description and also a discussion of the construction of likelihood for the underlying data. It also explains the notations used in the subsequent sections. In section 3, the different models are being proposed assuming various distributions for the lag times for estimation of the probability of perinatal HIV transmission and also to study of the suitability of the model is discussed. Application of proposed models to the data on the South African post exposure prophylaxis (PEP) has been discussed. An attempt is made to model the said transmission using exponential distribution, geometric distribution and shifted geometric distribution as lag time distribution. In section 4, we have discussed about the consistency of the parameters estimated and applicability of AIC and BIC for suitability of model selection and discussion of some related results. In section 5 we have discussed about the use of discrete distribution for modeling and compared the results with the case of continuous distribution for modeling. Section 6, is devoted to simulation study for the proposed models.

2. Modeling for perinatal HIV transmission

2.1. **Notations and related discussion.** In this section we revisit the analysis done by a model proposed by (N.Gupte et al., 2007) to estimate the probabilities of perinatal HIV transmission and the sensitivity of the viral assay used for HIV diagnosis in infants . Let η be the probability that the infection occurred either in antepartum or in intrapartum period. Further suppose that an infected infant is being breastfed at time t following the birth, let $\lambda(t)$ to be hazard rate of infection through breastfeeding at time t. Infants were tested for HIV infection using viral assays such as polymerase chain reaction(PCR). PCR is highly sensitive test that can detect small amounts of DNA or RNA (genetic material) in blood or tissue samples using an amplification techniques that multiplies the existing DNA/RNA so that it can be more easily detected. Let N be the number of mother infant pairs under study. Infants born to HIV infected mothers will be tested, using DNA PCR at the following ages (time points) $t_1, t_2, ... , t_k$ (k>1); where t_k is the visit after breastfeeding cessation. Let $\underline{X}_i = (X_{it_1}, X_{it_2}, ..., X_{it_k})$ denote results of DNA PCR assays at the scheduled visits $t_1, t_2, ..., t_k$ for i^{th} the infant; $i = 1, 2, ..., N$. Where,

$$X_{ij} = \begin{cases} 1 & ; \text{if DNA PCR is positive} \\ 0 & ; \text{if DNA PCR is negative} \end{cases} \tag{2.1}$$

$i = 1, 2, ..., N,\ j = t_1, t_2, ..., t_k$. Let d_i denote the duration of breastfeeding if the i^{th} infant is breastfed. If an infant is not breastfed then $d_i = 0$. Let ω denote the proportion of breastfed infants. The observed data are X $=(\underline{X}_1, \underline{X}_2, ..., \underline{X}_N)$ and D $=(d_1, d_2, ... d_N)$. It is assume that all HIV infected infants are identified using DNA PCR. It is also assumed that DNA PCR assay to be perfectly specific at all times. Unfortunately, the DNA PCR for detecting HIV infection is not perfectly sensitive. In fact, studies suggests that the sensitivity of PCR increases with an infants age (Owens, D.K et al., 1996). The sensitivity of PCR may depend on the timing of transmission and the time since infection. Our main interest is to study of sensitivity of PCR which is characterized by random variables U and V ; which are defined as follow U: Number of days after birth for getting a positive PCR with μ_u be the mean and $F_u(t)$ be distribution function of U. Further, V: Number of days after infection for getting a positive PCR; if the transmission takes place through breastfeeding with μ_v be the mean and $F_v(t)$ be distribution function of random variable V. Random variables U and V are considered as lag times for positive PCR assay. In view of above description one can observed that to model the maternal infant HIV transmission it is sufficient to model lag time distribution by an appropriate probability distribution. Unfortunately these lag time random variables are not observables hence usual procedure of direct estimation based on sample coming from the under lying distribution cannot be used. Hence it is proposed to assume various probability distributions as a model for lag

time distributions and carry out the analysis and models will be compared based on the nature of the estimates and also using Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). The first such attempt was made by Gupte et al with exponential distribution for the lag time distribution for random variables U and V. Since U and V are the number of days we thought that it is appropriate to model these using discrete distribution such as geometric distribution rather than exponential distribution. As mentioned above comparison among these models along with model given by N.Gupte et al., is carried out in subsequent section. The article also proposes some more models based on various lag time distributions such as geometric, shifted geometric for both U and V variables.

2.2. **The Likelihood.** In this section we revisit the likelihood developed by N.Gupte et al., for the estimation of $\underline{\theta}$ =(η, $\lambda(t)$, μ_u, μ_v). Here $\underline{\theta}$ is the parameter of interest. The likelihood function is given as

$$
\begin{aligned}
(2.2) \quad L(\underline{\theta}|X,D) &= [p_{t_1}^{BF}(\underline{\theta})]^{n_{t_1}^{BF}} * [p_{t_1}^{NBF}(\underline{\theta})]^{n_{t_1}^{NBF}} * \{\prod_{j=1}^{k-1}[p_{t_j,t_{j+1}}^{BF}(\underline{\theta})]^{n_{t_j,t_{j+1}}^{BF}}\} * \\
&\quad \{\prod_{j=1}^{k-1}[p_{t_j,t_{j+1}}^{NBF}(\underline{\theta})]^{n_{t_j,t_{j+1}}^{NBF}}\} * [p_{t_k}^{BF}(\underline{\theta})]^{n_{t_k}^{BF}} * [p_{t_k}^{NBF}(\underline{\theta})]^{n_{i_k}^{NBF}}
\end{aligned}
$$

where,

${n_{t_1}}^{BF}$: Number of breastfeeding infants who test positive at t_1,
${n_{t_1}}^{NBF}$: Number of non breastfeeding infants who test positive at t_1,
$n_{t_j,t_{j+1}}^{BF}$: Number of breastfeeding infants who test negative at time t_j and positive at time t_{j+1},
$n_{t_j,t_{j+1}}^{NBF}$: Number of non breastfeeding infants who test negative at time t_j and positive at time t_{j+1},
${n_{t_k}}^{BF}$: Number of breastfeeding infants who test negative at time t_k,
${n_{t_k}}^{NBF}$: Number of non breastfeeding infants who test negative at time t_k,

with

$$
\begin{aligned}
(2.3) \quad p_{t_1}^{BF}(\underline{\theta}) &= \omega * (1-\eta) * \int_0^{\min\{t_1,d_i\}} \lambda(u) * exp\{-\int_0^u \lambda(s)ds\} \\
&\quad * F_v(t_1-u)du + \omega * \eta * F_u(t_1)
\end{aligned}
$$

$$
(2.4) \quad p_{t_1}^{NBF}(\underline{\theta}) = (1-\omega) * \eta * F_u(t_1)
$$

$$
\begin{aligned}
(2.5)\quad p^{BF}_{t_j,t_{j+1}}(\underline{\theta}) &= \omega * (1-\eta) * \int_0^{\min\{t_j,d_i\}} \lambda(u) exp\{-\int_0^u \lambda(s)ds\} * \\
&\quad \{F_v(t_{j+1}-u) - F_v(t_j-u)\}du + \\
&\quad \int_{t_j}^{\min\{t_{j+1},d_i\}} \lambda(u) exp\{-\int_0^u \lambda(s)ds\} * F_v(t_{j+1}-u)du + \\
&\quad \eta * \omega * \{F_u(t_{j+1}) - F_u(t_j)\}
\end{aligned}
$$

$$
(2.6)\qquad p^{NBF}_{t_j,t_{j+1}}(\underline{\theta}) = (1-\omega) * \eta * (F_u(t_{j+1}) - F_u(t_j))
$$

$$
\begin{aligned}
(2.7) p^{BF}_{t_k}(\underline{\theta}) &= \omega * (1-\eta) * \int_0^{\min\{t_k,d_i\}} \lambda(u) exp\{-\int_0^u \lambda(s)ds\} \\
&\quad *(1 - F_v(t_k-u))du + \\
&\quad 1 - \int_0^{\min\{t_k,d_i\}} \lambda(u) exp\{-\int_0^u \lambda(s)ds\}du + \omega * \eta * \{1 - F_u(t_k)\}
\end{aligned}
$$

and

$$
(2.8)\qquad p^{NBF}_{t_k}(\underline{\theta}) = (1-\omega) * [\eta * (1 - F_u(t_k)) + (1-\eta)]
$$

respectively.

Observe that above likelihood is in the multinomial distribution functional form. From equations (2.3) to (2.8) the probabilities are the functions of the parameter $\underline{\theta}$ =(η, $\lambda(t)$, μ_u, μ_v) a vector valued parameter belonging to an admissible parameter set Θ. The true value of the parameter $\underline{\theta}^0$ supposed to be an interior point of parameter space Θ.(C. R. Rao,1973) .

3. **Proposed Models**

In the further discussion we proposed various models for HIV transmission based on various lag time distributions. As pointed out earlier these are based on the various distributional assumptions for the lag time random variables U and V. Initially we will discuss the form of likelihood under lag time distributions as an exponential distributions and results are revisited as those of N.Gupte el al.,2007. However strictly speaking the variables U and V can be regarded as discrete random variables hence it is most appropriate to assume some discrete distributions for random variables U and V, with this thought in mind it is decided to discuss the analysis by assuming U and V to be geometric random variables. We will also discuss in this section the case when shifted geometric distribution is used as a lag time distribution for r.v.s U and V. For the maximisation of likelihood and estimation of parameters purpose we have used the data on study made in South Africa as mentioned in N.Gupte el al.,2007.

This study data consists of one thousand and forty mother - infant pairs in the study, of which, 461 mothers breastfed their infants and 512 mothers chose an alternative feeding mechanism. For 67 women the feeding pattern was not available, so we used 973 women in our analysis. The post partum visits are taken at t_1= 3^{rd}day , t_2= 10^{th}day , t_3= 42^{nd}day , t_4= 90^{th} day , t_5= 180^{th} day. Based on these facts we proceed to take up modeling exercise and analysis for this study.

3.1. **Model-I: Exponential distribution.** The first proposed model is developed on the basis of the following assumptions (i) U and V have an exponential distribution with means μ_1 and μ_2 respectively (ii) hazard rate$\lambda(t)$ to be constant λ per month for the estimation of transmission rates.The likelihood $L(\underline{\theta}|X, D)$ for this model using equation (2.2) can be written as

$$p_{t_1}^{BF}(\underline{\theta}) = \omega * (1-\eta) \int_0^{t_1} \lambda * exp\{-\int_0^{u} \lambda ds\} * \{1 - e^{\frac{-(t_1-u)}{\mu_2}}\} du + \omega * \eta * \{1 - e^{\frac{-(t_1)}{\mu_1}}\} \tag{3.1}$$

$$p_{t_1}^{NBF}(\underline{\theta}) = (1-\omega) * \eta * \{1 - e^{\frac{-(t_1)}{\mu_1}}\} \tag{3.2}$$

$$\begin{aligned} p_{t_j,t_{j+1}}^{BF}(\underline{\theta}) = {} & \omega * (1-\eta) * \int_0^{t_j} \lambda * exp\{-\int_0^{u} \lambda ds\} * \\ & \{e^{\frac{-(t_j-u)}{\mu_2}} - e^{\frac{-(t_{j+1}-u)}{\mu_2}}\} du \\ & + \int_{t_j}^{t_{j+1}} \lambda * exp\{-\int_0^{u} \lambda ds\} * \{1 - e^{\frac{-(t_{j+1}-u)}{\mu_2}}\} du + \\ & \eta * \omega * \{e^{\frac{-(t_j)}{\mu_1}} - e^{\frac{-(t_{j+1})}{\mu_1}}\} \end{aligned} \tag{3.3}$$

$$p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}) = (1-\omega) * \eta * \{e^{\frac{-(t_j)}{\mu_1}} - e^{\frac{-(t_{j+1})}{\mu_1}}\} \tag{3.4}$$

$$\begin{aligned} p_{t_k}^{BF}(\underline{\theta}) = {} & \omega * (1-\eta) * \int_0^{t_k} \lambda * exp\{-\int_0^{u} \lambda ds\} * \{1 - e^{\frac{-(t_k-u)}{\mu_2}}\} du + \\ & 1 - \int_0^{t_k} \lambda * exp\{-\int_0^{u} \lambda ds\} du + \omega\eta\{1 - e^{\frac{-(t_k)}{\mu_1}}\} \end{aligned} \tag{3.5}$$

and

$$p_{t_k}^{NBF}(\underline{\theta}) = (1-\omega) * [\eta * (\{e^{\frac{-(t_k)}{\mu_1}}\}) + (1-\eta)] \tag{3.6}$$

respectively. The likelihood function $L(\underline{\theta}|X, D)$ was maximized numerically using Wolfram Mathematica 10.0 to obtain the estimate $\underline{\hat{\theta}}$. The estimates of the parameters obtained are $\hat{\eta}$ =0.1571138, $\hat{\lambda} = 0.000087520$, $\hat{\mu_1}$=10.4 days, $\hat{\mu_2}$=8.92 days. Using this model we also estimated the expected proportion of HIV - positive infants at each postpartum visit. Figure 1. compares the observed and the estimated proportion of HIV transmissions at each postpartum visit.

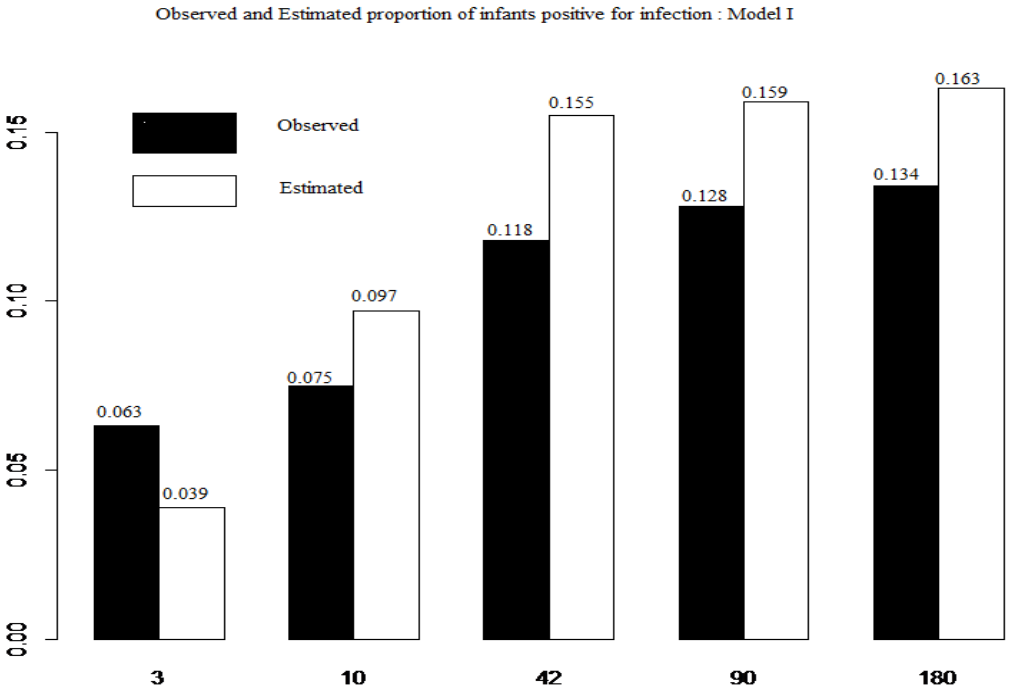

Figure 1.

Further we obtained AIC and BIC for Model I (refer Table 1.).
Now since the time points are days which are discrete , we thought that the lag times can be modeled by some discrete distribution. The analogue of exponential in discrete case is geometric. Hence the next model is proposed which has lag time distribution as geometric distribution.

3.2. **Model-II: Geometric Distribution.** In this case transmission rates were estimated on the basis of the following assumptions (i) U and V have an geometric distribution with parameters p_1 and p_2 respectively.(ii) the hazard rate $\lambda(t)$ to be constant equal to λ per month. The likelihood $L(\underline{\theta}|X, D)$ for this model can be obtained using equation (2.2) as

$$p_{t_1}^{BF}(\underline{\theta}) = \omega * (1-\eta) * \int_0^{t_1} \lambda * exp\{-\int_0^{u} \lambda ds\} * \{1-(q_2)^{(t_1-u)}\} du + \omega * \eta * \{1-(q_1)^{(t_1)}\} \tag{3.7}$$

$$p_{t_1}^{NBF}(\underline{\theta}) = (1-\omega) * \eta * \{1-(q_1)^{(t_1)}\} \tag{3.8}$$

$$
\begin{aligned}
(3.9)\qquad p^{BF}_{t_j,t_{j+1}}(\underline{\theta}) &= \omega*(1-\eta)*\int_0^{t_j}\lambda*exp\{-\int_0^{u}\lambda ds\}* \\
&\quad \{(q_2)^{(t_j-u)}-(q_2)^{(t_{j+1}-u)}\}du \\
&\quad +\int_{t_j}^{t_{j+1}}\lambda exp\{-\int_0^{u}\lambda ds\}*\{1-(q_2)^{(t_{j+1}-u)}\}du+ \\
&\quad \eta*\omega*\{(q_1)^{(t_j)}-(q_1)^{(t_{j+1})}\}
\end{aligned}
$$

$$
(3.10)\qquad p^{NBF}_{t_j,t_{j+1}}(\underline{\theta}) = (1-\omega)*\eta*\{(q_1)^{(t_j)}-(q_1)^{(t_{j+1})}\}
$$

$$
\begin{aligned}
(3.11)\, p^{BF}_{t_k}(\underline{\theta}) &= \omega*(1-\eta)*\int_0^{t_k}\lambda*exp\{-\int_0^{u}\lambda ds\}*\{1-(q_2)^{(t_k-u)}\}du+ \\
&\quad 1-\int_0^{t_k}\lambda*exp\{-\int_0^{u}\lambda ds\}du+\omega\eta\{1-(q_1)^{(t_k)}\}
\end{aligned}
$$

and

$$
(3.12)\qquad p^{NBF}_{t_k}(\underline{\theta}) = (1-\omega)*[\eta*((q_1)^{(t_k)})+(1-\eta)]
$$

respectively.

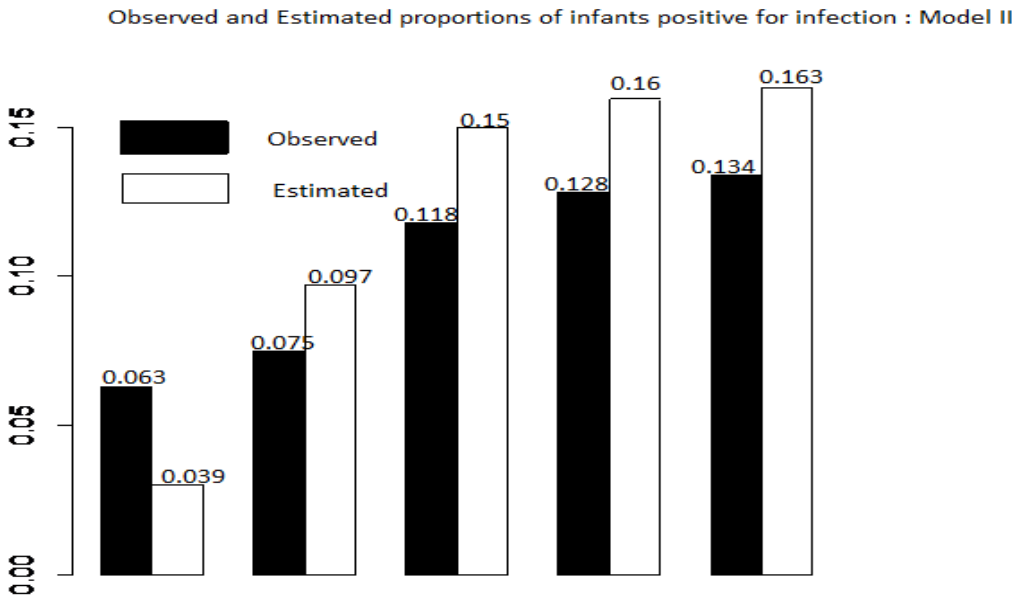

Figure 2.

The likelihood function $L(\underline{\theta}|X,D)$ was again maximized numerically using Wolfram Mathematica 10.0 to obtain the estimate $\hat{\underline{\theta}}$. The estimates of the parameters are $\hat{\eta}$ =0.15721172, $\hat{\lambda}$=0.11239378 , $\hat{p_1}$=0.09162033, $\hat{p_2}$=0.000087661034. Using this model we estimated the estimated proportion of HIV - positive infants at each post partum visit. Figure 2., compares observed and estimated proportion of infants positive for infection at different time points for Model II. Further the AIC, BIC for

Model -II are evaluated (refer Table 1.). Now we proposed a third model, where we assume the shifted geometric distribution as lag time distribution for r.v. U and V.

3.3. **Model-III: Shifted Geometric Distribution.** In this case transmission rates were estimated on the basis of the following assumptions (i) U and V have shifted geometric distribution with parameters $(\beta,\ p_1\ )$and $(\beta,\ p_2)$ respectively.(ii) the hazard rate to be constant λ per month. The probability mass function of shifted geometric distribution for variables U and V is given by

$$
\begin{aligned}
&P[U=1]=\beta\\
&P[U=k]=(1-\beta)*p_1*(1-p_1)^{(k-2)};\\
&k=2,3,..\\
&0<\beta<1, 0<p_1<1
\end{aligned}
\tag{3.13}
$$

$$
\begin{aligned}
&P[V=1]=\beta\\
&P[V=k]=(1-\beta)*p_2*(1-p_2)^{(k-2)};\\
&k=2,3,...\\
&0<\beta<1, 0<p_2<1
\end{aligned}
\tag{3.14}
$$

For this Model , the likelihood $L(\underline{\theta}|X,D)$ for this model can be given as in equation (2.2) with

$$
\begin{aligned}
p_{t_1}^{BF}(\underline{\theta}) &= \omega*(1-\eta)*\int_0^{t_1}\lambda*exp\{-\int_0^{u}\lambda ds\}*\\
&\quad \{\beta+(1-\beta)*(1-(q_2)^{(t_1-u)})\}du+\\
&\quad \omega*\eta*\{\beta+(1-\beta)*(1-(q_1)^{(t_1)})\}
\end{aligned}
\tag{3.15}
$$

$$
p_{t_1}^{NBF}(\underline{\theta}) = (1-\omega)*\eta*(\beta+(1-\beta)*(1-(q_1)^{(t_1)}))
\tag{3.16}
$$

$$
\begin{aligned}
p_{t_j,t_{j+1}}^{BF}(\underline{\theta}) &= \omega*(1-\eta)*\int_0^{t_j}\lambda*exp\{-\int_0^{u}\lambda ds\}*\\
&\quad \{(1-\beta)*((q_2)^{(t_j-u)}-(q_2)^{(t_{j+1}-u)})\}du+\\
&\quad \int_{t_j}^{t_{j+1}}\lambda*exp\{-\int_0^{u}\lambda ds\}*\\
&\quad \{\beta+(1-\beta)*\{1-(q_2)^{(t_{j+1}-u)}\}\}du+\\
&\quad \eta*\omega*(1-\beta)*\{(q_1)^{(t_j)}-(q_1)^{(t_{j+1})}\}
\end{aligned}
\tag{3.17}
$$

$$
p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}) = (1-\omega)*\eta*(1-\beta)*\{(q_1)^{(t_j)}-(q_1)^{(t_{j+1})}\}
\tag{3.18}
$$

$$
\begin{aligned}
(3.19)\quad p_{t_k}^{BF}(\underline{\theta}) &= \omega * (1-\eta) * \int_0^{t_k} \lambda * exp\{-\int_0^{u} \lambda ds\} * \\
&\quad (1-\beta) * \{(q_2)^{(t_k-u)}\} du + \\
&\quad 1 - \int_0^{t_k} \lambda * exp\{-\int_0^{u} \lambda ds\} du + \omega * \eta * (1-\beta)\{(q_1)^{(t_k)}\}
\end{aligned}
$$

and

$$
(3.20)\qquad p_{t_k}^{NBF}(\underline{\theta}) = (1-\omega) * [\eta * \{(1-\beta) * (q_1)^{(t_k)}\} + (1-\eta)]
$$

respectively. The likelihood function $L(\underline{\theta}|X, D)$ was maximized numerically using Wolfram Mathematica 10.0 to obtain the estimate $\hat{\underline{\theta}}$. The parameters estimate are : $\hat{\eta}$=0.15559921, $\hat{\lambda}$=0.00001096, $\hat{\beta}$=0.32229, $\hat{p_1}$=0.04725571, $\hat{p_2}$=0.003408526. Using this model we estimated the expected proportion of HIV - positive infants at each post partum visit. Figure 3. compares the observed and the expected proportion of HIV transmissions at each postpartum visit. Further the AIC, BIC for Model -III are evaluated (refer Table 1.). Now we proceed to make discussions abut results and the

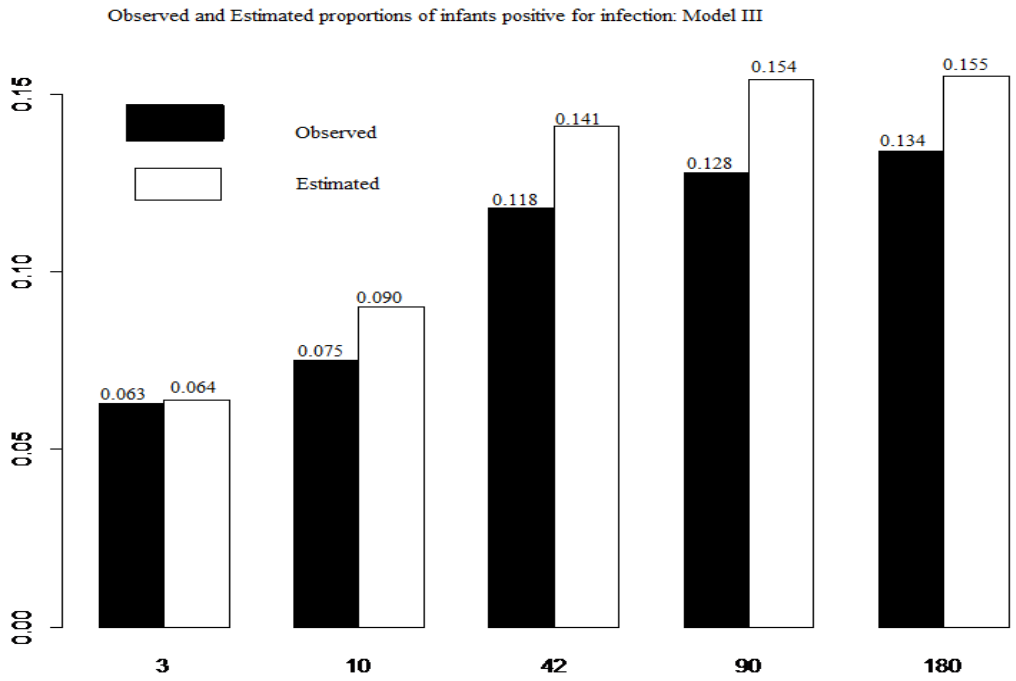

Figure 3

comparative study of the models in view of average difference error in observed and estimated proportion of infants as well as in terms of Akaike information criterion (AIC) and Bayesian information criterion(BIC).

4. **Discussion**

Our attempt towards modeling maternal infant HIV transmission by considering different lag time distributions for the variables U and V gave fruitful results. N. Gupte et al. already discussed about the same by taking exponential distribution as lag time distribution. But we have made an attempt to justify the estimation of parameters of interest, in terms of consistent estimator. Not only that we also considered the discrete distributions as lag time distribution, which we thought as suitable distributions for the given situation. In our set up we have assumed that the transmission rate is constant, which may not be the case over the different time periods. It is shown that under the following assumptions the likelihood admits consistent and unique solution . Consistent in a sense that the solution converges in probability to the true value $\underline{\theta}^0$.

Assumption 1 :Given $\delta > 0$, it is possible to find an ϵ such that

$$\{[p_{t_1}^{BF}(\underline{\theta}^0)] * log\frac{[p_{t_1}^{BF}(\underline{\theta}^0)]}{[p_{t_1}^{BF}(\underline{\theta})]}\} + \{[p_{t_1}^{NBF}(\underline{\theta}^0)] * log\frac{[p_{t_1}^{NBF}(\underline{\theta}^0)]}{[p_{t_1}^{NBF}(\underline{\theta})]}\}$$
$$+\{\sum_{j=1}^{k-1}[p_{t_j,t_{j+1}}^{BF}(\underline{\theta}^0)] * log\frac{[p_{t_j,t_{j+1}}^{BF}(\underline{\theta}^0)]}{[p_{t_j,t_{j+1}}^{BF}(\underline{\theta})]}\} +$$
$$\{\sum_{j=1}^{k-1}[p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}^0)] * log\frac{[p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}^0)]}{[p_{t_j,t_{j+1}}^{NBF}(\underline{\theta})]}\}$$
$$+\{[p_{t_k}^{BF}(\underline{\theta}^0)] * log\frac{[p_{t_k}^{BF}(\underline{\theta}^0)]}{[p_{t_k}^{BF}(\underline{\theta})]}\} + \{[p_{t_k}^{NBF}(\underline{\theta}^0)] * log\frac{[p_{t_k}^{NBF}(\underline{\theta}^0)]}{[p_{t_k}^{NBF}(\underline{\theta})]}\}\} \geq \epsilon$$

where $|\underline{\theta}\text{-}\underline{\theta}^0|$ is the distance between $\underline{\theta}$ and $\underline{\theta}^0$.

Assumption 2 :

$$p_{t_1}^{BF}(\underline{\theta}) \neq p_{t_1}^{BF}(\underline{\beta})$$
$$p_{t_1}^{NBF}(\underline{\theta}) \neq [p_{t_1}^{NBF}(\underline{\beta})]$$
$$p_{t_j,t_{j+1}}^{BF}(\underline{\theta}) \neq [p_{t_j,t_{j+1}}^{BF}(\underline{\beta})]; j = 1, 2, ...k-1$$
$$p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}) \neq [p_{t_j,t_{j+1}}^{NBF}(\underline{\beta})]; j = 1, 2, ...k-1$$
$$p_{t_k}^{BF}(\underline{\theta}) \neq [p_{t_k}^{BF}(\underline{\beta})]$$
$$p_{t_k}^{NBF}(\underline{\theta}) \neq [p_{t_k}^{NBF}(\underline{\beta})]$$
$$when \underline{\theta} \neq \underline{\beta}$$

Assumption 3: The functions $p_{t_1}^{BF}(\underline{\theta}), p_{t_1}^{NBF}(\underline{\theta}), p_{t_j,t_{j+1}}^{BF}(\underline{\theta}), p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}), p_{t_k}^{BF}(\underline{\theta}), p_{t_k}^{NBF}(\underline{\theta})$are continuous in $\underline{\theta}$.

Assumption 4: The functions $p_{t_1}^{BF}(\underline{\theta}), p_{t_1}^{NBF}(\underline{\theta}), p_{t_j,t_{j+1}}^{BF}(\underline{\theta}), p_{t_j,t_{j+1}}^{NBF}(\underline{\theta}), p_{t_k}^{BF}(\underline{\theta}), p_{t_k}^{NBF}(\underline{\theta})$ admit first order partial derivatives

Assumption 5 : The functions in assumption 4 are continuous at $\underline{\theta}^0$

Assumption 6: The information matrix (i_{rs}) be non singular at $\underline{\theta}^0$ where

$$\begin{aligned} i_{rs} &= \frac{1}{p^{BF}_{t_1}(\underline{\theta})}\frac{\partial p^{BF}_{t_1}(\underline{\theta})}{\partial\theta_r}\frac{\partial p^{BF}_{t_1}(\underline{\theta})}{\partial\theta_s}+\frac{1}{p^{NBF}_{t_1}(\underline{\theta})}\frac{\partial p^{NBF}_{t_1}(\underline{\theta})}{\partial\theta_r}\frac{\partial p^{NBF}_{t_1}(\underline{\theta})}{\partial\theta_s}+ \\ &\sum_{j=1}^{k-1}\frac{1}{p^{BF}_{t_j,t_{j+1}}(\underline{\theta})}\frac{\partial p^{BF}_{t_j,t_{j+1}}(\underline{\theta})}{\partial\theta_r}\frac{\partial p^{BF}_{t_j,t_{j+1}}(\underline{\theta})}{\partial\theta_s}+\sum_{j=1}^{k-1}\frac{1}{p^{NBF}_{t_j,t_{j+1}}(\underline{\theta})}\frac{\partial p^{NBF}_{t_j,t_{j+1}}(\underline{\theta})}{\partial\theta_s}\frac{\partial p^{NBF}_{t_j,t_{j+1}}(\underline{\theta})}{\partial\theta_s}+ \\ &\frac{1}{p^{BF}_{t_k}(\underline{\theta})}\frac{\partial p^{BF}_{t_k}(\underline{\theta})}{\partial\theta_r}\frac{\partial p^{BF}_{t_k}(\underline{\theta})}{\partial\theta_s}+\frac{1}{p^{NBF}_{t_k}(\underline{\theta})}\frac{\partial p^{NBF}_{t_k}(\underline{\theta})}{\partial\theta_r}\frac{\partial p^{NBF}_{t_k}(\underline{\theta})}{\partial\theta_s} \end{aligned}$$

Assumption 1 and 4 imply that an maximum likelihood estimator can be obtained as a root of the equations $\frac{\partial L(\underline{\theta}|X,D)}{\partial\theta_r}$=0, r=1,2,3,4 with probability 1 and is unique. Assumptions 1, 5, 6 imply that maximum likelihood estimator is a consistent solution of the likelihood equation. These results are generalisations of Cramer - Huzurbazar theory. It is verified that the Assumptions 1-6 are satisfied using Wolfram Mathematika 10.0 (details not shown) that shows that the maximum likelihood estimators (m.l.e) obtained are consistent solution of the likelihood equation.We show as a sample case these conditions are satisfied in case of exponential lag times as well as geometric lag times. This ensures a unique solution to a likelihood and it is also a consistent and asymptoticlly normal providing Best Asymptotically normal estimator (BAN) for $\underline{\theta}$. Note that this fact is not discussed in the case of exponential lag times by (N. Gupte et al.,2007). Hence the above discussion is add on to the work done by these authors and also provides justification for the analysis carried out by them. The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Hence, AIC provides a means for model selection.

5. **Comparative study of the Models**

Average difference error in observed and estimated proportion of infants positive for infection for each Model is evaluated. It is noticed that the average difference error in observed and estimated proportion for Model I and Model II are close enough. It also indicates as the models proposed (Model I and Model II) are performing as same as in case of Model I given by N.Gupte et al. (2007). But it is interesting to note that the selection intelligent choice of lag times for random variables U and V as shifted geometric distribution serves as the model with the interesting results. It is observed that the average difference error in observed and estimated proportion for Model III is certainly less at every stage of testing of DNA/ PCR assay as compared to Model I and Model II. From this fact Model III can be considered as one of the best model for modeling Maternal Infant HIV transmission. It is also reflected in

TABLE 1. Table showing AIC , BIC for Model-I,II,III.

Model	AIC	BIC	AIC Corrected
Model-I	2378.2	2397.71	2378.24
Model-II	2376.5	2396.01	2376.04
Model-III	2345.28	2345.24	2369.67

in terms of AIC / BIC. We have computed the AIC , BIC and Corrected AIC for all the models (refer Table 1.). It is observed that the model II performs better by AIC criterion. However the difference between AIC and BIC are small for the models hence further improvement if possible is recommended. But the values of AIC and BIC for Model III is the least as compared to all other Models.On the basis of AIC , Model III is the best for performance of the analysis. The results are encouraging as also noticed in terms of observed and estimated proportion of infants positive for infection.

6. **Simulation Study**

We have carried out the simulation study for Model-I ,by performing 1000 simulations each with 1000 observations by taking different values of the parameters of the variables U and V. For the variable U we have considered mean values μ_1 to be 9.9 , 10.0 , 10.1 and 10.2 and for variable V we have considered mean values μ_2 to be 7.9 , 8.0 , 8.1 , 8.2 . Using Wolfram Mathematica 10 the simulation is carried out and the estimates of the parameter η is obtained along with empirical standard deviation of then estimated parameterl.Further distribution Fit test is also carried out and almost all cases the null hypothesis that the data is distributed according to the Normal distribution is not rejected at the 5 percent level based on the Cramer - Von Mises test. The results are encouraging. We have also carried out the simulation study for Model-III ,by performing 1000 simulations each with 1000 observations by taking different values of the parameters of the variables U and V. For the variable U we have considered values p_1 to be 0.04 , 0.03 , 0.02 and 0.01 and for variable V we have considered values p_2 to be 0.05 , 0.04 , 0.03 , 0.02 . Using Wolfram Mathematica 10 the simulation is carried out and the estimates of the parameter η is obtained along with empirical standard deviation of the estimated parameter and it is tabulated in Table 2.

TABLE 2. Results of simulation study : Shifted Geometric distribution.

$\downarrow Parameters \rightarrow$	$p_1 = 0.04$	$p_1 = 0.03$	$p_1 = 0.02$	$p_1 = 0.01$
$p_2 = 0.05$	$\eta = 0.160(0.0005)$	$\eta = 0.160(0.0005)$	$\eta = 0.160(0.0005)$	$\eta = 0.160(0.0005)$
$p_2 = 0.04$	$\eta = 0.165(0.0007)$	$\eta = 0.165(0.0007)$	$\eta = 0.165(0.0007)$	$\eta = 0.165(0.0007)$
$p_2 = 0.03$	$\eta = 0.173(0.0010)$	$\eta = 0.173(0.0001)$	$\eta = 0.173(0.0010)$	$\eta = 0.173(0.001)$
$p_2 = 0.02$	$\eta = 0.190(0.001)$	$\eta = 0.190(0.0015)$	$\eta = 0.192(0.0015)$	$\eta = 0.190(0.0015)$

REFERENCES

[1] A review of available evidence (2004). *WHO Library Cataloguing - in- Publication Data HIV transmission through breastfeeding A review of available evidence (2004)*

[2] C.R.Rao(1973).*Linear Statistical Inference and its Applications , 2^{nd} Edition, Wiley- inter science*

[3] Conor, E.M. Sperling R., et al (1994). Reduction of maternal infant transmission of human immunodeficiency virus type 1 with Zidovudine treatment. *The New England Journal of Medicine 331, 1172-1180.*

[4] Owens, D.K. Holodniy, M. et al (1996). A meta - analytic evavluation of the polymerses chain reaction for thr diagnosis of HIV infection in infants. *Journal of American Medical Association 275 , 1342-1348*

[5] Paolo,G.M , Taha, E.T, et al.(1999). HIV transmission through breastfeeding, A study in Malawi.*Journal of American Medical Association 282-2 , 744-749*

[6] N.Gupte , R. Brookmeyer , et al (2007). Modeling Maternal - Infant HIV Transmission in the presence of breastfeeding with an imperfect test. *Biometrics 63 , 1189-1197*

[7] Newton L.k. ,Hoover D.R , et al (2008). Extended Antiretroviral Prophylaxis to reduce breast-milk HIV-1 transmission. *The New England Journal of Medicine 359 , 119-129*

[8] Raji Balasubramanian and Lagakos S.W. (2001). Estimation of timing of perinatal transmission of HIV.*Biometrics 57 , 1048-1058*

[9] Raji Balasubramanian and Lagakos S.W. (2003). Estimation of a failure time distribution based on imperfect diagnistic tests.*Biometrika 9 , 171-182*