A SPECIAL "SURVIVAL" DISTRIBUTION PROBABILITY FUNCTION TO DESCRIBE NEW-BORNS RELATED TO THE HIV/AIDS INFECTION

H. M. Yang, M. Della Negra, Y. C. Lian, W. Queiroz & L. K. Hotta
A SPECIAL "SURVIVAL" DISTRIBUTION PROBABILITY FUNCTION TO DESCRIBE NEW-BORNS RELATED TO THE HIV/AIDS INFECTION

H. M. Yang*, M. Dalla Negra*, Y. C. Lim*, W. Queiroz* & L. K. Hotta*

*UNICAMP / IMECC - Departamento de Matemática Aplicada
Instituto de Infectologia Emílio Ribas
**UNICAMP / IMECC - Departamento de Estatística

SUMMARY

As well as adults, children may become infected by HIV, most frequently by the so-called vertical or perinatal transmission. A statistical approach is proposed to deal with the seroconversion (the cleansing of maternally derived antibodies against HIV by seropositive children) and survival of the vertically HIV infected children. The model is fitted to followed up seropositive children data. We estimate by maximum likelihood both half-age and heterogeneity degree, and compare them with non-parametric Kaplan-Meier estimation with respect to the half-age.

INTRODUCTION

The infection of children with Human Immunodeficiency Virus (HIV), by the so-called vertical (or perinatal) transmission route, has been reported since 1982 (Anderson and May, 1992). Since then the epidemiological evidence (Connor, 1987) pointed to a prenatal acquisition of HIV appears to be the principal route of transmission. In Brazil, this route of transmission is responsible for over 92% of HIV infected children, because the most important risk factor for HIV infection among female population is through sexual relationship (Ministério da Saúde, 1995). Other routes may occur as perinatal transmission in utero via transplacental infection of the fetus, during labour and delivery by contact with infected blood and secretions, and postnatally via breast milk (CDC, 1994). When an anti-HIV antibodies screening of the at-risk new-borns (from HIV infected parturients) are performed, not all the seropositive children are HIV infected because some of them may harbour only maternally derived antibodies against HIV. Nevertheless, the following-up of these seropositive new-borns, accompanied by the anti-HIV test, can be a useful epidemiological tool. This kind of study permits us to observe two mutually exclusive reachable situations by a seropositive newborn. The first, seroconversion phenomenon, is characterized by the continuous decaying of maternally derived antibodies against HIV. This decaying can be approximated by an exponential function (Azevedo, 1994). The second, the survival, is determined by the periods of time that vertically infected HIV children can survive even when there is a cumulative destruction of the helper/inducer
subpopulations of T-lymphocytes by HIV, which leads to the deficiency in the immunity response. This immuno-deficiency is the gate by which other infections can be initiated and rapidly spread out, leading sometimes to the death. As the previous study, we assume that the destruction of T-lymphocytes follows an exponential function.

Our task is the development of a methodology to deal with the followed-up seropositive children from the Emilio Ribas Infectology Institute, located in São Paulo State, Brazil. The data consists of 583 seropositive new-borns followed-up from December, 1985 until September, 17th 1996, reporting the age at which they seroreverted or died. Note that both phenomena are concerned to physical and laboratorial observations, which are the level of antibodies anti-HIV concentration (seroreversion) and the remaining amount of T-lymphocytes cells (survival). The values at which we can assure the occurrence of the seroreversion or death due to HIV/AIDS are greatly varied among the followed up children, therefore, the data record related to both phenomena are in terms of the age at which occurred seroreversion or death. We develop a simple statistical model, based on two parameters, named half-age (the age at which the value of the probability of being in one of the state is one half) and heterogeneity degree (different responses of children under external stimuli), to analyze the transition from the initial (seropositive or survival) to final (seronegative or death) states.

A STATISTICAL APPROACH

We determine the probability of leaving the initial state based on physical (clinical, laboratorial, etc.) observations regarded to the health condition. The statistical model for the transition time is related to a latent variable, say $W$, which is based on health condition of the child. Lower values of $W$ indicate the maintaining of the initial state. For instance, the surviving of child infected with HIV is related to a higher concentration of CD4 cells that indicates a better health condition, while high levels of maternally derived antibodies anti-HIV concentration in uninfected child is related to seropositive state.

Observe that the latent variable $W$ relates inversely with the health condition. Additionally, let us consider that the distribution of this latent variable at the transition is supposed to be exponential (Cox, 1992) with hazard rate $\lambda$, i.e., with the probability density function given by

$$f_W(w) = \lambda e^{-\lambda w}, \lambda > 0; \quad w > 0.$$ 

However the variable health condition $W$ is not necessarily observable. For instance, it is well understood that AIDS symptoms, and consequently the death by an opportunistic disease, is related to the lowering in the amount of T-lymphocyte cells, which is, in general, not measured. In the same fashion, the following up of the HIV seropositive children consists of performing anti-HIV test at each 6 months. Conditionally to the health condition, the distribution of the age is proportional to the distribution of $\log(X+1)$, where $X$ has exponential distribution with (constant) hazard rate $W$, according to

$$f_X(x) = we^{-wx}, \quad w > 0; \quad x > 0.$$
The rationale behind it is that the expectation of the transition age, conditionally to the health condition, decreases proportionally to the logarithm of the health condition $w$. Now, calling as $a$ the age at the transition, we have

$$f_{aw}(a|w) = \frac{w}{a} \xi e^{-\frac{1}{\xi}(e^{\frac{w}{a}} - 1)} , \quad w > 0 ; \quad a > 0 ,$$

where $\xi$ is an auxiliary parameter relating age and physical observation.

Thus, the probability that the child keeps the initial state, conditionally to $W=w$, is given by

$$S_{aw}(a|w) = e^{-\frac{w}{a}(e^{\frac{w}{a}} - 1)} , \quad w > 0 ; \quad a > 0 .$$

But, we are taking the physical observation (health condition) as a weight, hence the probability that a child keeps the initial state before or at age $a$ is given by

$$S(a) = \left[ 1 + \lambda \left( e^{\frac{a}{\beta}} - 1 \right) \right]^{-1} ,$$

where $\beta = \frac{1}{\xi}$ is the heterogeneity degree. We can relate the parameter $\lambda$ with the half-age by the expression $\lambda = e^{\frac{a}{\beta}} - 1$, where $\xi$ is the age at which we have $S(\xi) = 1/2$.

Finally, we can obtain the probability density of occurring the transition from the initial to final state at age $a$ by

$$s(a) = e^{\frac{a}{\beta}} \left[ \frac{1}{\beta} + \frac{e^{\frac{a}{\beta}} - 1}{e^{\frac{a}{\beta}} - 1} \right]^{-2} ,$$

which comes from the derivative relation $s(a) = -dS(a)/da$.

**RESULTS AND CONCLUSIONS**

For the estimation of parameters to maintain the seropositive state (seroreversion study), we considered 27 seropositive children as censored and 172 seronegative children as transited, in a total of 199 (102 females and 97 males) children. For the estimation of parameters related to keeping the surviving state with respect to vertically HIV infected children (survival study), we considered 27 seropositive and 192 HIV and AIDS symptomatic children as censored (partial sum 219) and 170 HIV/AIDS children who died as transited, in a total of 389 (184 females and 205 males) children.

Firstly, these followed-up children from Emilio Ribas Infectology Institute are described by non-parametric Kaplan-Meier estimation. With respect to half-age, we found 11.10 months for the seroreversion and 91.23 months for the survival. If we take into account the gender, the estimated half-age, for seroreversion, are 11.03 and 11.44 months, respectively for females and males, while for survival, 73.81 and 101.31 months, respectively for females and males.

In Figures 1 and 2 we show the results obtained from the non-parametric Kaplan-Meier estimation for both seropositive state maintaining and survival of HIV/AIDS infected children.
Figure 1. The seropositive state maintaining probability function estimated by the Kaplan-Meier.

Figure 2. The survival probability function estimated by the Kaplan-Meier.
With respect to the parametric approach, we estimate both the half-age and heterogeneity degree by applying the maximum likelihood estimation (Yang et al., 1997, Yang and Coutinho, 1998). In order to do this, we maximize the logarithm of the likelihood function, \( l \), defined by

\[
l = \sum_{i} \ln[s(a_{i})] + \sum_{c} \ln[S(a_{c})],
\]

where \( S(a) \) and \( s(a) \) are, respectively, probability and probability density functions given previously, and the subscripts \( i \) and \( c \) stand, respectively, for children who transitioned from initial state to final state and who are censored.

With respect to half-age, we found 12.01 months for the seroreversion and 81.16 months for the survival. If we take into account the gender, the estimated half-age, for seroreversion, are 12.07 and 11.87 months, respectively for females and males, while for survival, 67.60 and 98.03 months, respectively for females and males. With respect to heterogeneity degree, we found 4.10 months for the seroreversion and 220.8 months for the survival. If we take into account the gender, the estimated heterogeneity degree, for seroreversion, are 4.31 and 3.88 months, respectively for females and males, while for survival, 89.6 and 1252.8 months, respectively for females and males.

In Figures 3.a and 3.b we show the results obtained from the parametric estimation for the seropositive state maintaining.

![Graph showing seropositive probability function estimated by the parametric estimation.](image)

Figure 3.a. The seropositive state maintaining probability function estimated by the parametric estimation.
Figure 3.b. The corresponding (seropositive) transition probability function estimated by the parametric estimation.

In Figures 4.a and 4.b we show the results obtained from the parametric estimation for the survival of HIV/AIDS infected children.

Figure 4.a. The survival probability function estimated by the parametric method.
Figure 4.b. The corresponding (survival) probability density function estimated by the parametric method.

The early seroreversion of males in relation to females shown by parametrized estimation can be better understood by looking the estimated heterogeneity degree. Observe that the male children (3.88) are more homogeneous than females (4.31). This parameter justifies the reversion in the seroreversion age when the parametrized estimation is compared with non-parametrized one. But, in relation to survival, we observe great inhomogeneity among males. Based on the estimated parameters, female children, in comparison with males, seem to have slow decaying of maternally derived antibodies, but less resistance against HIV infection, which leads to quick death due to HIV/AIDS.

Now, we compare the parametrized model with Kaplan-Meier estimation method. The results provided by the latter method was taken as the gold standard values. Observe that in Kaplan-Meier estimation, the data must be ordered in ascending age, and therefore, the position of the point in the sample is important: the former points contribute much more to the estimation than the latters. Also, the value of the probability is varied when the status of the individual is modified (transition occurred), while the censored individuals retain the previously calculated value of the probability. Observe that the Kaplan-Meier estimation presents higher surviving ages than the parametric approach. In this estimation, due to the assumption that the censored children are under the same future experience of death, the surviving HIV/AIDS children pull up the surviving curve. However, our parametrized approach considers to censored children a probability of being alive before or at the censored age, while those who died are described by a probability density which describes the exact age at death. For this reason, Kaplan-Meier estimation provides with higher value than our method. Finally, even though the model is very simple
with only two parameters, it fits nicely to the data for both phenomena (Yang et al., 1998a & 1998b).

REFERENCES