International Journal of Mathematics and Physical Sciences Research ISSN 2348-5736 (Online)
Vol. 5, Issue 2, pp: (52-57), Month: October 2017 - March 2018, Available at: www.researchpublish.com

Calculation of the Transition Probabilities and Intensities of HIV Progression in Kenya

Stephen Kinyanjui Mwambura[1], Mwangi Daniel Wambugu[2]

[1,2]Department of Mathematics and Actuarial Science, Catholic University of Eastern Africa

Abstract: **The methods that were employed in this paper analyzed HIV data. The aim was to evaluate the evolution of HIV positive patients to bring out some significant factors associated with this pathology on individuals in Kiambu county, Kenya. Many clinical situations can be described in terms of the conditions that individuals can be in, how they can move among such states (transitions), and how likely such moves are (transition probabilities). State transition models were monitored by serial measurements of CD4 cells per unit volume (mm^3) of blood in an HIV patient as a tool for modeling HIV disease progression. WHO disease staging system for HIV infection was used to analyze data. HIV progression was analyzed through the application of a four state Markov model with reversible transitions. Some of the major findings of the study were that patients presenting with a CD4 count between 200 and 349 had a far lower chance of immune recovery and a significant chance of immunosuppression compared to patients of a higher CD4 count. Transition probabilities from states 1, 2, and 3 into state 4 increased as time progressed. The estimated total length of stay in state 1 was longer than state 2 and 3 respectively. The rate of decline in CD4 count decreased at lower levels of the indicator.**

Keywords: **transition probabilities, transition intensities, HIV progression, CD4 count, Markov model.**

1. INTRODUCTION

The staged Markov model is a useful way of describing a process in which an individual move through a series of states in continuous time. Survival analysis is the simplest two-state model where individuals remain alive until an observed or censored time of death. Homogeneous Markov models are the most appropriate models since the relationship between transition probabilities and transition intensities are easy to determine. It is often useful for a model to include the duration of time spent in a certain state before it transits to another state. When considering different staged models, there are various related tools that were used to evaluate the data. Likelihoods, intensities or hazards, transition probabilities, CD4 counts, and distributions are some of the tools that were used to model the data discussed in this project. This study has analyzed HIV progression using CD4 counts intervals of six months since enrolment on ART with the objective of investigating the probabilities of transitions to lower CD4 counts and estimating the average stay in the CD4 count states.

2. STATEMENT OF THE PROBLEM

Standard survival analysis techniques like Kaplan Meier survival curves and Cox regression models can be used when one wishes to study the time until a specified event occurs. However, this does not describe a sequence of events and fails to utilize all the information that longitudinal data can provide if multiple events are studied. This study discusses the developments of homogeneous Markov models in HIV progression. One may examine the occurrence of clinical events which occur early in HIV infection or the values of biological markers as markers of disease progression. The most widely used marker of HIV progression is the CD4 lymphocyte count, which plays a crucial role in the immune system and is used to determine when ART should commence. When an individual loses CD4 cells he or she is more vulnerable to opportunistic infections. In this project, a model with state structure based on intervals of CD4 count and death as an absorbing state was used.

3. LITERATURE REVIEW

The desirability of a descriptive tool for censored survival data, free of parametric assumptions, had been recognized for decades. During the 1950s, well established demographic and actuarial techniques were presented to the medical-statistical community in influential surveys such as those by Berkson and Gage (1952). The penetrating study of Kaplan and Meier (1958), the fascinating genesis of which was chronicled by Breslow et al (1991), eliminated in principle the need for these approximations in the common situations in medical statistics where all survival and censoring times are known precisely.The homogenous Markov staged models are time-continuous models for which the transition probability only depends upon the current state and has constant transition hazards for the state. In medical research, discrete observation times are generally used. The state that the patient is in at the observation is the only thing known. The researcher may know that there has been a transition to a new state but does not know when in that interval or time it occurred. Thus we consider homogeneous Markov models with interval censoring.

In 2011, Tarylee used the *msm* package to study the application of multistate Markov models to HIV disease progression in South Africa. She used a model with state structure based on four intervals of CD4 count and ARV initiation an absorbing state. The model was successfully fitted to data from the "Sinikithemba" cohort, consisting of 451 HIV positive individuals residing in Durban, South Africa. Reverse transitions were incorporated into the model. The rate of CD4 count decline was examined and predictions of the trajectory of patients were made through the use of transition probability matrices. The likelihood is calculated from the transition probability matrix $P(t)$. For a time-homogeneous process, the (r, s) entry of $P(t)$, $p_{rs}(t)$, is the probability of being in state s at a time $t + u$ in the future, given the state at time u is r. It does not say anything about the time of transition from r to s, indeed the process may have entered other states between times u and $t + u$. $P(t)$ can be calculated by taking the matrix exponential of the scaled transition intensity matrix (Cox and Miller,1965)

$P(t) = \text{Exp}(t\ Q)$

The matrix exponential Exp is different from a scalar exponential. The exponential of a matrix is defined by the same "power series" $\text{Exp}(X) = 1 + X^2/2! + X^3/3! + \ldots$ as the scalar exponential, except that each term X^k in the series is defined by matrix products, not element-wise scalar multiplication. For simpler models, it is feasible to calculate an analytic expression for each element of $P(t)$ in terms of Q.

A model is governed by a transition intensity matrix Q. With an assumption that the patient can advance or recover from consecutive states while alive, and die from any state, the model with four states is given by

$$Q = \begin{pmatrix} -(q_{12} + q_{14}) & q_{12} & 0 & q_{14} \\ q_{21} & -(q_{21} + q_{23} + q_{24}) & q_{23} & q_{24} \\ 0 & q_{32} & -(q_{32} + q_{34}) & q_{34} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

4. RESEARCH METHODOLOGY

The conditional probability of entering state j at time t, given the system was in state i at time 0 is given by

$$p_{ij}(t) = P(X(t) = j \mid X(0) = i) \text{ for all } i, j \in S \text{ and } t \geq 0$$

The quantities $p_{ij}(t)$ are called the transition probabilities, and form a matrix $P(t)$ called the transition probability matrix. Hence for a CTMC with n possible states we can define

$$P(t) = \begin{pmatrix} p_{11}(t) & \cdots & p_{1n}(t) \\ \vdots & \ddots & \vdots \\ p_{n1}(t) & \cdots & p_{nn}(t) \end{pmatrix}$$

We are now able to describe the behavior of a CTMC using the exponential distribution property.

Suppose that once a system enters state *i*at time *t* it remains there for an exponentially distributed period of time with parameterq_i, where $q_i \geq 0$, and then jumps to state *j* ($j \neq i$). It will be noted that despite the assumption of time homogeneity, the state specific times are assumed to have separate exponential distribution rates thereby allowing state heterogeneity.

5. FINDINGS

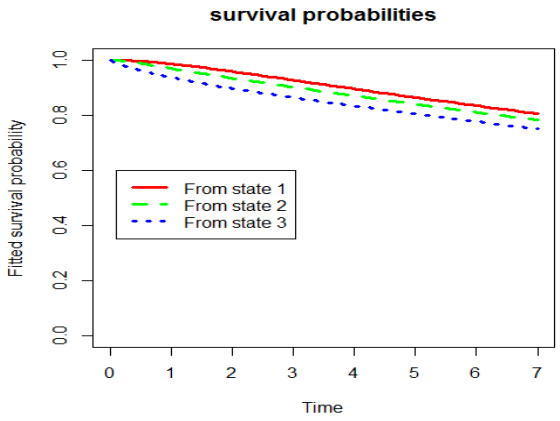

Figure 1.Plot of the expected probability of survival versus time from transient states

In figure 1, Survival is defined as not entering the absorbing state. This shows that a four year survival advanced immunosuppression is approximately 0.86 as opposed to 0.90 with mild and 0.92 without a significant immunosuppression.

The transition intensity matrix is given by

$$Q = \begin{bmatrix} -1.055 & 1.049 & 0 & 0.005744 \\ 1.088 & -2.452 & 1.354 & 0.009982 \\ 0 & 1.484 & -1.584 & 0.1003 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

These are calculated as $\frac{-1}{\hat{q}_{ii}}$ where $\hat{q}_{ii}$ is the i^{th} diagonal entry of the estimated transition intensity matrix Q.

The transition probability matrix $P(t)$ is evaluated below at times 1, 2, 3, 4, 5, 6, and 7 years. As time progressed the probability for transitions to the absorbing state increased and the probability of immune recovery decreases.

$$P(1) = \begin{bmatrix} 0.5114 & 0.3032 & 0.171 & 0.01434 \\ 0.3146 & 0.3494 & 0.305 & 0.03098 \\ 0.1945 & 0.3343 & 0.4075 & 0.06373 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(2) = \begin{bmatrix} 0.3902 & 0.3182 & 0.2496 & 0.04196 \\ 0.3301 & 0.3194 & 0.2847 & 0.06576 \\ 0.2839 & 0.312 & 0.3013 & 0.1029 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(3) = \begin{bmatrix} 0.3492 & 0.3129 & 0.2655 & 0.07333 \\ 0.3247 & 0.3069 & 0.2699 & 0.09854 \\ 0.3019 & 0.2958 & 0.2665 & 0.1358 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(4) = \begin{bmatrix} 0.3282 & 0.3037 & 0.2632 & 0.1049 \\ 0.3151 & 0.2959 & 0.2591 & 0.1299 \\ 0.2993 & 0.284 & 0.2505 & 0.1663 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(5) = \begin{bmatrix} 0.3146 & 0.2936 & 0.256 & 0.1358 \\ 0.3046 & 0.2855 & 0.2497 & 0.1601 \\ 0.2911 & 0.2737 & 0.2399 & 0.1953 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(6) = \begin{bmatrix} 0.303 & 0.2835 & 0.2477 & 0.1658 \\ 0.2942 & 0.2756 & 0.241 & 0.1892 \\ 0.2816 & 0.2641 & 0.231 & 0.2233 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P(7) = \begin{bmatrix} 0.2923 & 0.2737 & 0.2392 & 0.1947 \\ 0.284 & 0.266 & 0.2326 & 0.2173 \\ 0.272 & 0.2549 & 0.229 & 0.2502 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Interpretation of the transition probability matrix

Transition probabilities from states 1, 2, 3 into state 4 (absorbing state) increases as time increases. For example

$p_{24}(1)$ = 0.03098 while $p_{24}(7)$ = 0.2173

$p_{34}(1)$ = 0.06373 while $p_{34}(7)$ = 0.2502

A patient presenting with 200⩽ CD4 count<349 has a 25% chance of being absorbed within seven years.

6. CONCLUSION AND RECOMMENDATION

Based on the findings of the study the following conclusions were made at 95% confidence limits: The study reveals that a patient presenting with 200⩽ CD4 count<349 has a 25% chance (0.1966, 0.4042) of being absorbed within seven years. It is also clear that transition probabilities from states 1, 2, and 3 increases as time increases. For example $p_{24}(1)$ = 0.03098 (0.02297, 0.07481) while $p_{24}(7)$ = 0.2173 (0.1683, 0.3987).This model has been applied to one health facility and in one county. Since counties in Kenya are unique, it will be important to apply the model using datasets from each county for comparison of effects of covariates on process parameters accurately if state transitions are observed exactly.

REFERENCES

[1] Aalen,O.O.&Johansen,S.(1978).An *empirical transition matrix form on homogeneous Markov chains based on censored observations*.Scandinavian Journal of Statistics 5, 141–150.

[2] Armitage, P. and Doll, R. (1954). *The age distribution of cancer and a multistage theory of carcinogenesis*. Br. J. Cancer, 8, 1-15. United States of America.

[3] Beck, J.R., Pauker, S.G (1983). *The Markov process in medical prognosis*. Med Decision Making 3,419-458.

[4] Berkson and Gage (1952). *TheSimplest cure rate model.* Wikipedia. The Free Encyclopedia.

[5] Binquet, C., Rey, D., Mahboubi,A. (2009). *Markov modelling of HIV infection evolution in Highly Active Antiretroviral Therapy*. France.

[6] Breslow K. & McLaren (1991). *Thick Malignant Melanomas with good clinical outcome: a histological study and survival analysis*. East of Scotland.

[7] Bwayo, J., Nagelkerke, N., Moses, S., Embree, J., Ngugi, E., Mwatha, A., Kimani, J., Anzala, A., Choudhri, S., and Achola, J. (1995). *Comparison of the declines in CD4 counts in HIV-1-seropositive female sex workers and women*

from the general population in Nairobi, Kenya. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology10, 457–461.

[8] Christopher Jackson (2011). *Multi-state modeling with R: the msm package*. Journal of Statistical Software Volume 38, Issue 8, 1-25. Medical Research Council Biostatistics Unit, Cambridge, U.K.

[9] Cox, R.D & Miller, H.D (1965). *The Theory of Stochastic Processes*. Chapman and Hall, London.

[10] Cox,D.R.(1972). *Regression models and life-tables* (with discussion). Journal of the Royal Statistical Society Series B 34, 187-220.

[11] Ira M. Longini, Jr, Michael G.Hudgens (2003). *Lecture notes on stochastic processes in biostatistics: Applications to infectious diseases*. Pp 31-61. Atlanta

[12] Jensen, P.A., Bard, J.F., 2002b. *Sensitivity analysis, duality and interior point methods. In: Operations Research Models and Methods*, John Wiley & Sons, Inc, Danvers, MA pp.111-145

[13] Kaplan, E.L & Meier, P. (1958). *Nonparametric estimation from incomplete observations.* Journal of the American Statistical Association 53(282), 457-481.

[14] Kay, R. (1986). *A Markov model for analyzing Cancer Markers and Disease States in Survival Studies*. Biometrics 42:4, 855-865.

[15] Longini, I. M., Clark, W. S., Byers, R. H., Ward, J. W., Darrow, W. W., George F.L, and Herbert W.H. (1989). *Statistical analysis of the stages of HIV infection using a Markov model. Statistics in Medicine*8, 831–843.

[16] Longini, I. M., Clark, W. S., Gardner, L. I., and Brundage, J. F. (1991). *The Dynamics of CD4+ T-Lymphocyte decline in HIV-Infected Individuals: A Markov Modeling Approach*. Journal of Acquired Immune Deficiency Syndromes 4, 1141–1147.

[17] Meeker, W.Q., Escobar, L.A. (1998b). *Models for continuous failure time processes. In Statistical Methods for Reliability Data*, John Wiley and Sons, Inc, New York pp.26-45.

[18] Mullins, D. C. (1996). *A Simplified Approach to Teaching Markov Models*. American Journal of pharmaceutical Education 60, 42–47.

[19] Redfield, R.R., Wright D. and Tramont,E(1986). The *Walter Reed staging classification for HTLV-III/LAV infection*. New England Journal of Medicine, 314, 131-132

[20] Simpson K.N, Baran R.W (2011). *Markov models to assess the cost-effectiveness of switching to a second-line protease inhibitor regimen of lopinavir plus ritonavir for HIV-infected patients who failed a first-line treatment in sub-Saharan Africa.* University of New York.

[21] Simwa, R.O. &MugishaJ .T,(2003c). A *stochastic model for the CD4 cells counts in an HIV/AIDS patient and its application in treatment interventions*. In: Proceedings of the 1st Pan-African Biomathematics Congress 8th-12th, December, 2003, Kampala, Uganda.

[22] Simwa, R.O.,(2003a). *On empirical modeling of HIV/AIDS pandemic with application to East Africa*. African J.Sci. Technol. 4, 34-43.

[23] Statistics (2007). *Random processes*. University of KwaZulu-Natal, Durban, South Africa.

[24] Tarylee, R. (2011). *The Application of Multistate Markov Models to HIV Disease Progression*, University of KwaZulu-Natal,MSc thesis, South Africa.

[25] Titman,A.C (2007). *Model diagnostics in multi-state models of biological systems.* PhD thesis, University of Cambridge.

[26] World Health Organization (2005). *Interim Clinical Staging of HIV/AIDS & HIV/AIDS case definitions for surveillance*, African Region.