A STOCHASTIC ANALYSIS OF THE ABSORPTION PROBABILITIES OF CD4 CELL COUNTS OF HIV/AIDS PATIENTS USING THE SMOOTHED NON-STATIONARY MARKOV CHAIN MODEL. A CASE STUDY OF ANAMBRA STATE

Osisiogu U. A.

Department Of Industrial Mathematics and Computer Science Ebonyi State University, Abakaliki, Ebonyi State, Nigeria

Nwosu C. A.

Department Of Statistics Anambra State University, Uli, Anambra State, Nigeria

ABSTRACT: The question of when an HIV patient should be subjected to therapy using highly active antiretroviral therapy (HAART) is very challenging. The direct application of the existing probabilistic models, like the stationary Markov Chain for the determination of the life expectancies of patients and their absorption probabilities might not always be reliable. In this work, we propose the smoothed non-stationary Markov chain approach which is conceptually efficient than the stationary Markov chain and non-stationary Markov chain models. We examined a total of 1094 HIV patients (cohort) from January - December 2012 with follow-up in their CD₄ cell transition counts, collected from the Medical Examination Department of the Nnamdi Azikiwe University Teaching Hospital, Continuous Quality Improvement HIV Care Unit (NAUTH), Nnewi – Anambra State – Nigeria. The patients were grouped into five immunological states developed by Guiseppe Di Biase et al (2007). The five states considered were as follows: state one ($CD_4 > 500 \text{ cells/mm}^3$), state two $(350 < CD_4 \le 500 \text{ cells/mm}^3)$, state three $(200 < CD_4 \le 350 \text{ cells/mm}^3)$, state four $(CD_4 \le 350 \text{ cells/mm}^3)$ > 200 cells/mm³), state five (Death). These states define the seriousness of the sickness based on the epidemiological state of the patients CD₄ cell counts. The estimation of the smoothed non-stationary probabilities were computed using the exponential smoothing technique. The results obtained show that patients in state I, state II, State III, and state IV, have the following absorption probabilities: 0.6159, 0.6080, 0.5915 and 0.6122. The results also show that the absorption probabilities of patients with low CD₄ counts do not differ appreciably from patients with higher CD₄ cell counts, meaning that low CD₄ cell counts do not generally imply a faster rate of absorption of patients suffering from the infection [Osisiogu U. A., Nwosu C. A (2013)], studies have shown that patient's age is a relevant factor to the rate of absorption.

KEYWORDS: Smoothed Non-stationary Markov Chain Model, Exponential Smoothing Technique, Absorption Probability, CD₄ cell counts, Immunological States.

INRODUCTION

The outbreak of the HV/AIDS epidemic in Nigeria in 1981 and 1983 respectively and the report of the cases of the disease in the thirty-six states of the federation including Abuja, the Federal capital territory made researchers to carry out different studies on the disease; like the determination of the life expectancy of patients [Osisiogu U. A. and Nwosu C. A (2013)], predicting future CD₄ cell counts of HIV patients [Osisiogu U. A. and Nwosu C. A (2013)], comparing models to determine the most efficient one that can predict when to start the

highly active antiretroviral therapy (HAART) [Osisiogu U. A. and Nwosu C. A. (2013)]. The study provides the use of epidemiological data in estimating the impact of HIV/AIDS dynamics using the Smoothed Non-stationary Markov Chain Model. The HIV fatal effect arises from the attack of the CD₄ cell counts which play a pivotal regulatory role in the immune response to infections and tumours [Anderson R. M. et al (1986)]. The hallmark of the HIV/AIDS infection is the death of the patient. A number of approaches have been used for the HIV/AIDS dynamics, we shall demonstrate the use of the Smoothed Non-stationary Markov Chain Model and the fundamental matrix of the absorbing Markov Chain in this study.

NOTATIONS

i T = Calendar time in months, T = (0, 1, 2, ...)

ii K = the number of states in the system.

 $iii \qquad n_{ij}(T) = \text{the number of patients in state } i \text{ at month } T, \text{ who transited to state } j \text{ at month } T+I.$

iv $n_i(T)$ = the number of patients in state i at month T.

v $n_{i,k}(T)$ = the number of patients in state i who died at month T

vi $P_{ii} = n_{ii}(T)$

 $n_i(T)$: Probability of patients in state i at month T, who

transited to state j at the end of the month T.

vii $W_i = n_{i,k}(T)$

 $n_i(T)$: death rate of patients in state i at month T.

viii $N(T) = \sum_{n=1}^{N} (T)$: total number of patients in state I at the beginning of the month T.

ix $n_{0j}(T+I)$: New entrants into state j at the beginning of the month T.

THE MARKOVIAN MODEL FOR THE CD₄ CELL COUNTS FOR THE SMOOTHED NON-STATIONARY MARKOV CHAIN MODEL.

Let the difference equation:

$$\overline{n_{j}}(T+I) = \sum_{i} \mathbf{n}_{i,j}(T) + n_{o,j}(T+I) \dots [1]$$

$$(i = 1, 2, ... N) (T+0, 1, 2, ...)$$

denote the expected values of the CD_4 cell counts, where the bars are the expected values (i = 1, 2, ... k). The above equation in other words is saying that patients in state j are patients who transited to state j. Some of these variables may assume zero values especially when the study has to do with a cohort, where no newly infected patients are allowed in, i.e.

$$n_{o,c} = 0$$
 where $c = Cohort$ study.

We can express the transition flow of the newly infected patients as wastages i.e.

$$n_{o,i}(T) = W_i(T+I)$$

Then equation [1] becomes

$$\frac{1}{n_j}(T+I) = \sum_{\substack{i=j\\i=j}} \mathbf{n}_{ij}(T) + W_j(T+I) \dots [2]$$

Step 1

$$P_{ij} = \sum n_{ij}(T)$$

$$\sum n_{i}(T)$$

$$T=I$$

$$T=I$$

$$T=I$$

Substituting equation [3] in equation [2]

$$\overline{n_i}(T+I) = \sum n_i(T) P_{ij} + W_i(T+I) ...$$
[4]

In vector form, equation [4] becomes:

$$\overline{n}(T+I) = n(T) P + W(T+I) \dots [5]$$

Step 2

If the probability is non-stationary, the probability is denoted by notation [6]:

$$P_{ij}(T) = n_{ij}(T) \qquad \qquad ---- \qquad \qquad [6]$$

Substituting equation [6] into equation [2], we have

In vector form, equation [7] becomes:

$$\overline{n}(T+I) = n(T) P(T) + W(T+I)$$
 [8]

Step 3

The Smoothed Non-stationary Markov Chain Model is an extension of the Non-stationary Markov Chain Model, whose structure is known by the T- step transition probability matrix, unlike the Stationary Markov Chain that is associated with powers of one step transition probability matrix P, which is a common estimate of the transition probability matrices over the past months on the assumption that they are stationary over time [Osisiogu U. A. (2004)]. We note that these estimates are the sum of the estimates for each month, where the weights are proportional to n(T). If we vary these weights and put more weights on the current transition matrices and less weights on the previous, we obtain the Smoothed Non-stationary Markov Chain Model, and this depends on the stochastic variation prevalent at the time of the data, Equation [6] now becomes:

$$P_{ij}(T) = \sum_{i=0}^{N} \beta_{i}(r) P_{ij}(T) \dots [9]$$

Where
$$r = [T^* - T], [T = 0, 1, 2, ..., T^*]$$

 $\beta_i(r) = \alpha_i (I - \alpha_i)^r$ [10]

We can now rewrite equation [9] as:

$$R_{ij}(T) = \alpha_i P_{ij}(T) + (I - \alpha_i) P_{ij}(T - I) \dots [11]$$

$$(i, j = 1, 2, \dots k) (T = 0, 1, 2, \dots T^*)$$

and P_{ij} (T*) is as defined in equation [9]. Thus, the transition probability matrix P used in this model for future prediction is the one whose elements are derived from equation [11] given.

that:

In vector form, the smoothed Non-stationary Markov Chain Model becomes:

$$\overline{nj}(T+I) = n(T) P(T) + W(T+I) \dots [13]$$

٨

4

Published by European Centre for Research Training and Development UK (www.eajournals.org)

CANONICAL FORM

A Markov Chain with absorbing states can be represented in a canonical form by renumbering the states such that the transient states come first. If there are r absorbing states and t transient states, the transition matrix will have the following canonical form:

$$P = \begin{pmatrix} Q & R \\ \dots & Q \\ O & I \end{pmatrix}$$
 [14]

Where I is an rxr identity matrix, O is an rxt zero matrix, R is a nonzero txr matrix and O is a txt matrix. The first t states are transient and the last r states are absorbing.

FUNDAMENTAL MATRIX

For an absorbing Markov Chain P, the matrix $N = (I - Q)^{-1}$ is called the fundamental matrix for P. The entry n_{ii} of N gives the expected number of times the HIV patients are in the transient state i having started from the transient state i. The fundamental matrix N helps in the calculating of the life expectancy of these patients in state j who entered the system in state i, and also to determine the absorption probabilities of these patients assuming that these variables are stochastic random variables prevailing in the HIV/AIDS dynamics. We can establish the existence of the inverse of the matrix (I - O) from the following definition and theorem [Kendall M. G. and Stuart A. (1961)].

DEFINITION

Let X be the length of time unit t, a patient spends in state j, starting initially from state i. Let U = E(X) be the expected length of time unit t a patient spends in state j, starting initially from state i ii

THEOREM

Let Q be the transition matrix of the infected HIV patient. Let [U] = U^t , where $[U^t_{ij}]$ is defined above. Then: $U^{t} = (I_{ij} Q)^{-1}$ Where U = [U] is a matrix. $t - \infty$ [15]

PROOF

Let Z be an indicator function defined as:

after t time unit

1, If a patient in state i moves to state j
$$Z_{ij}^{t} = 0, \quad \text{Otherwise}$$

The probability distribution of Z for a fixed t is given as:

$$P[Z = 0] = I - Q \qquad [16]$$

$$ij \qquad ij$$

$$ij \qquad ij$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$P[Z = I] = I - Q$$
[17]

Then the expected value of Z is given as:

$$E[Z] = Q \cdot_{t} \dots I \cdot_{ij} \dots [18]$$

$$= (I - Q^{T}) (I - Q)^{-1}$$

Where I is an identity matrix. Given that
$$U^t = [U]$$
 then $U^t = (I_{\overline{Lin}}Q)^{-1}$[20]

Equation [20] is called the fundamental matrix of P.

ABSORPTION PROBABILITIES

THEOREM

Let b_{ij} be the probability that a patient will be absorbed in the absorbing state j, having started initially from the non-absorbing state i. Let B be the matrix with entries b_{ij} . Then B is a txr matrix and

$$B = NR [21]$$

Where N is the fundamental matrix and R is as in the canonical form.

PROOF

We have

$$\begin{split} B_{ij} &= \Sigma \Sigma q \quad r_{kj} \\ &= \Sigma \Sigma q \quad r_{kj} \quad \stackrel{(n)}{\underset{ikj}{n}} \quad \stackrel{(n)}{\underset{ikj}{l}} \\ &= \Sigma^k r_{ik}^n \stackrel{il_{(n)}}{\underset{ikj}{r_{kj}}} \\ &= (N R)_{ij} \end{split}$$

7. APPLICATION

To illustrate the efficiency of the models, we apply it to a cohort study of 1094 HIV patients with follow-up in their transition counts from Jan – Dec 2012.

DATA

The data were sourced from the medical examination department of the Nnamdi Azikiwe University Teaching Hospital, Continuous Quality Improvement HIV Care (NAUTH) Nnewi, Anambra State – Nigeria.

METHOD

The CD₄ cell counts of these patients were classified into five immunological states based on the classification developed by Guiseppe Di Biase et al (2007). The five different states are represented by ranges of CD₄ cell counts:

State I:
$$CD_4 > 500 \text{ cells/mm}^3$$

State II: $350 < CD_4 \le 500/mm^3$

State III: $200 < CD_4 \le 350/\text{mm}^3$ State IV: $CD_4 > 500 \text{ cells/mm}^3$ State V: Death (Absorbing State)

The transition probability matrices of the cohort study for the 12 observable months were recorded thus:

Table I: nij for the month of January

JANUARY		I	II	III	IV	V	TOTAL
$CD_4 > 500 \text{ cells/mm}^3$	I	95	99	48	18	12	272
$350 < CD_4 \le 500/mm^3$	II	100	96	50	20	4	270
$200 < CD_4 \le 350 / mm^3$	III	70	50	95	55	7	277
$CD_4 > 500 \text{ cells/mm}^3$	IV	76	95	30	64	10	275
Death	V	0	0	0	0	0	0
Total							1094

Table 1: Represents the transition counts of the CD4 cells of the 1094 patients for the month of January 2012, using the above classification. Similar classifications were done for the months of Feb – Dec. 2012.

<u>Table II: Transition Probability Matrix for the Month of Jan 2012.</u>

		₹	II	III	IV	V
	I	0.349	0.364	0.176	0.066	0.044
	II	0.370	0.356	0.185	0.074	0.015
P =	III	0.252	0.181	0.343	0.199	0.025
	IV	0.276	0.354	0.109	0.233	0.036
	V	0	0	0	0	1

The transition probability matrix for the month of January in table II was computed using equation (6) and similar computations were done for the months of February – December 2012. By equation (II) and a smoothing constant, $\alpha = 0.30$, we obtain the transition probability matrix $P = [P_{ij}]$ whose elements are used in obtaining the life expectancy and absorption probabilities of the patients.

^

Table III: Transition Probability Matrix P for the Month of Jan – Dec. 2012.

$$I = \begin{pmatrix} I & II & III & IV & V \\ 0.330 & 0.327 & 0.206 & 0.116 & 0.021 \\ II & 0.318 & 0.341 & 0.205 & 0.112 & 0.015 \\ 0.237 & 0.246 & 0.318 & 0.150 & 0.016 \\ IV & 0.242 & 0.325 & 0.256 & 0.157 & 0.018 \\ V & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

By equation [14], we represent the transition probability matrix P in canonical form.

Table IV: Showing the Canonical Form of the Transition Probability Matrix P.

$$I = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0.330 & 0.327 & 0.206 & 0.116 & 0.021 \\ 0.318 & 0.341 & 0.205 & 0.112 & 0.015 \\ 0.237 & 0.246 & 0.318 & 0.150 & 0.016 \\ 1V & 0.242 & 0.325 & 0.256 & 0.157 & 0.018 \\ V & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Table IV represents the canonical form of the transition probability matrix P, where Q, R, O, and I are shown as in the canonical form.

$$O = [0\ 0\ 0\ 0]$$
 $I = [1]$

From the canonical form, the matrix (I - Q) is obtained and the inverse, $N = (I = Q)^{-1}$ computed (see equation 20).

Table V: Showing the Life Expectancy of Patients.

I
$$(1.275 \quad 10.941 \quad 8.382 \quad 4.497$$

II $(10.232 \quad 11.922 \quad 8.356 \quad 4.479$

N = $(I - Q)^{-1}$ = III $(1.506 \quad 9.248 \quad 4.396)$

IV $(1.171 \quad 10.928 \quad 8.436 \quad 5.539)$

The total life expectancy of patients before absorption are obtained by:

$$\Sigma u_{ii} = NC$$

Where Σ u_{ij} is the total Tife expectancy, N, the fundamental matrix and C, a column vector all of whose entries are I.

From the equation [21], we compute the absorption probability of the patients. From the canonical form

$$R = \begin{array}{c} I & \left(0.021 \\ II & 0.015 \\ \end{array} \right)$$

$$IV & \left(0.018 \right)$$

$$B = NR \qquad \text{where } N = \text{fundamental matrix}.$$

$$I \qquad 0.6159$$

$$II \qquad 0.6080$$

$$0.5915$$

$$IV \qquad 0.6122$$

B denotes the absorption probabilities in each state of the Markov Chain.

RESULT AND DISCUSSION

The results obtained show that patients in state I, state II, state III and state IV have the following absorption probabilities: 0.6159, 0.6080, 0.5915, and 0.6122. Some studies carried out on HIV/AIDS dynamics have shown that patient's age is a relevant factor to forecast the transitions among the different levels of seriousness of the disease [Zelalem Getahun Dessie (2014)]. The probability that an HIV/AIDS patient in any of the good states will transit to the absorbing state (death state) is greater with increasing age, irrespective of the current state and age of the patient. More generally, the probability of being absorbed decreases with increasing CD₄ cell counts over time. Therefore, if patients are not subjected to therapy at the appropriate time using the highly active antiretrovirals (HAART), their absorption probabilities increases as a result of the decrease in their CD₄ cell counts [Osisiogu U. A. and Nwosu C. A. (2013)]. The role of the CD₄ cell count in HIV management cannot be overemphasized and the impact of HAART over the past 20 years has made HIV more of a chronic disease for practitioners to manage, requiring careful clinical monitoring. Laboratory markers such as the HIV-1 RNA viral load and CD₄ cell count are regularly used for patient management in addition to predicting disease progression and/or treatment outcomes. The HIV viral load is considered to be the gold standard for evaluating treatment success and absorption probabilities, although it is often limited to cost. The CD₄ cells are vital utility immunological components for the prediction of HIV disease progression and time for absorption of any HIV patient. The articulation of these variables: the CD₄ cell / HIV-1 RNA viral loads will aid clinicians to examine the added value of the CD₄ cell counts in the management of a person with HIV infection.

CONCLUSION

The Smoothed Non-stationary Markov Chain Model is applied to capture the HIV/AIDS dynamic progression and to determine the absorption probabilities of HIV patients. The model considers the length of stay (life expectancy) of the HIV patients, the randomness in the different states in which the infection can evolve and the probability of these patients being absorbed. The following can be concluded from this study.

The probability of dying decreases with increasing CD₄ cell counts over time. At any time of the process, there is more likely to be in worse state than be in a better one. In general, the absorption probability of an HIV/AIDS patient depends on his/her current state of the disease in such a way that lower CD₄ cell counts are associated with high risk of being absorbed. The dynamic nature of the AIDS progression is confirmed with particular findings that there is more likely to be in worse state than better one unless interventions are made. It is recommended that patients should be advised to keep up the ongoing HAART treatment services in most effective ways with the careful considerations of recent disease status of the HIV/AIDS patients.

REFERENCES

- Osisiogu, U. A. and Nwosu, C. A. (2013): Comparison between Stationary Markov Chain Model and Non-stationary Markov Chain Model in the tracking of CD₄ Cell Counts in an HIV/AIDS Patients and Determining their Life Expectancy. International Journal of Mathematical Science (IJMS) ICASTOR India Vol. 7(2) (2013), PP 95 108.
- U. A. Osisiogu, Nwosu, C. A. (2013): Comparison between Stationary Markov Chain Model and Non-stationary Markov Chain Model in Predicting CD₄ Cell Counts in an HIV/AIDS Patients. International Journal of Mathematical Science (IJMS) ICASTOR India Vol. 7 No. 2 2013), PP 153 – 168.
- Anderson, R. M., White, P. K. (1986): The consequence of HIV if not treated. Journal of Statistics and Computer Science, Vol. 11 (3), PP 332 336.
- Bartholomew, D. J. (1963): Stochastic Models for Social Process. 2nd Ed. Wiley London.
- Osisiogu, U. A. (2004): Application of Non-stationary Markov Chain Model to a Graded Manpower System in Oyo State Nigeria. Journal of Science Vol. 38, PP 35 51.
- Osisiogu, U. A. (2005): Application of Probabilistic Method in the Analysis of Career Prospect of Staff in Oyo State Civil Service Structure. Journal of Information Communication and Computing Technologies Vol. 1(1), PP 81 87.
- Guiseppe Di Baise, Gughelimo D' Amico, Arturo Di Girolomo, Jacque Jansee, Stefano Jacobili, Nicola Tinari, Raimando Manca (2007): A Stochastic Model for HIV/AIDS Dynamic Evolution. Journal of Mathematical Problem in Engineering Vol. 1, article 65636.
- U. A. Osisiogu, Nwosu, C. A. (2015): Predicting Future CD₄ Cell Counts of HIV/AIDS Patients by Non-stationary Markov Chain. A Case Study of Anambra State. Journal of Data Science Vol. 13(1) 140204.
- Anderson, R. M., Medley, G. F., May, R. M., Johnson, A.M. (1986): A Preliminary Study of the Transmission Dynamics of the Human Immunodeficiency Virus (HIV), the Causative Agent of AIDS. IMA Journal of Applied Medical Biology Vol. 3(4), PP 229 263.
- Cousine, M. S., Shickle, L. M., Bander, J. A. (2002): An Introduction to Predictive Modelling of Disease Management Risk Stratification. Journal of Disease Management Vol. 2(1), PP. 157 165.
- Annah, M. Jeffrey, Xiahua Xia (2003): When to Initiate HIV Therapy. A Control Theoritic Approach. IEEE Transaction on Biomedical Engineering Vol. 50, No. 11.
- Satten, G. A., Longini, I. M. (1999): Markov Chain with Measurement Error. Estimating the "True" Course of a Marker of the Progression of Human Immunodeficiency Virus Disease. Journal of Applied Statistics Vol. 45(3), PP 275 309.

- Goshu, A. T., Dessie, Z. G. (2013): Modelling Progression of HIV/AIDS Disease Stages Using Semi-Markov Processes. Journal of Data Science Vol. 11, PP 269 280.
- Zelalem Getahun Dessie (2014): Modeling of HIV/AIDS dynamic evolution using non-homogeneous Semi-Markov Process. Springer Pius. A Springer Open Journal. http://www.springerpius.com/content/3/1/537.