Journal of the Nigerian Association of Mathematical Physics
***Volume 25, No. 2* (November, 2013), pp 363 – 368**
© *J. of NAMP*

Application of Stationary Markov Chain Model in Predicting Future CD_4 Counts of HIV/AIDS Patients: A Case Study of Anambra State Nigeria

Nwosu C.A.

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

Abstract

We present a stationary Markov chain model for the prediction of future CD_4 cell counts of HIV/AIDS patients before HAART. Recorded hospital data were obtained for a cohort of 1004 patients with follow-up from the medical Examination department of Nnamdi Azikiwe Uniersity Teaching Hospital, continuous Quality Improvement HIV Care (NAUTH) Nnewi-Anambra Nigeria from January-December 2010. States of the Markov chain model are defined by the seriousness of the sickness based on the epidemiological counts of the CD_4 cell/mm^3. The five states considered are: State one (CD_4 cell counts > 500cells/mm^3), State two ($350 < CD_4 < 500$cells/mm^3), State three ($200 < CD_4 \leq 350$ cells/mm^3) State four ($CD_4 < 200$ cells/mm^3) and state five (Death). The first four states are named good or alive State. The result obtained from the mean absolute error (MAE) – 15.63 showed the stationary Markov chain model to be conceptually a good model in the prediction of future CD_4 cell counts.

Keywords: Stationary Markov chain, Transition Probability, Transition probability matrix, CD_4 cell counts.

1.0 Introduction

The human immunodeficiency virus (HIV) is among the most pressing health problem in the world today. It has been estimated that about 35.5 million people were living with the disease as at the end of 2008, ad greater proportion of the population coming from Africa and Asian countries.

The HIV fatal effect stems from the attack on a persons CD_4 cell counts. This result to the progressive depletion of the CD_4 cell counts which play a pivotal regulatory role on the immune response to infections and tumours [1]. Infection by the human immunodeficiency Virus (HIV) gradually evolves to acquire immunodeficiency syndrome (AIDS) and which finally leads to death. The use of Markov chain model [1 – 12] in predicting the future CD_4 cell counts of the HIV/AIDS patients is based on the fact that essential characteristics of Markovian model (i.e. (i) a finite set of states (ii) successive intervals and (iii) movement among states expressed in terms of probabilities) have direct analogies with the HIV/AIDS dynamics. States can be defined as CD_4 cells. Transition probabilities can defined as the rate of flow between CD_4 cell counts during fixed time interval. More so, the possibility matrix make it very attractive. The matrix of the forecasting model may help physician, policy makers, young researchers and innovators.

2.0 The Markovian Model for CD_4 Counts of the Stationary Markov Chain Model

The expected number of CD_4 cell counts in any state is related to the equation:

$$\bar{n_j}(T+1) = \sum_{i=1}^{N} n_{ij}(T) + n_{0.j}(T+1), (i = 1,2 \ldots\ldots k)(T = 0,1,2,\ldots). \quad (2.1)$$

Where the bar denote the expectation (j = 1,2,…,k) [2,3]. The above equation implies that patients in state j are patients who transited to state j. some of these variables may assume zero value, like in the present study which is a cohort and no new entrants are allowed, in this case;

$$n_{0.c}(T+1) = 0, where\ c = cohort\ study$$

Corresponding author: E-mail: vikkycee@yahoo.com, Tel.: +2348134214959

However, the above model could be redefined by expressing the transition flow of the new entrants in terms of wastage i.e., death rate of patients.

$n_{0.j}(T+1) = W_j(T+1)$

The equation (2.1) becomes

$$n_j(T+1) = \sum_{i=1}^{N} n_{ij}(T) + n_{0.j}(T+1), (i = 1,2 \ldots\ldots k)(T = 0,1,2, \ldots). \quad (2.2)$$

And the common estimate of the transition probability given as;

$$P_{i.j} = \frac{\sum_{T=1}^{T^*} n_{i.j}(T)}{\sum_{T=1}^{T^*} n_{i.j}(T)}, (i,j = 1,2 \ldots, K) \quad (2.3)$$

On the assumption that they are stationary over time. Then equation (2.1) becomes

$$n_j(T+1) = \sum_{i=1}^{N} P_{i.j} n_i(T) + W_j(T+1), (i = 1,2,3 \ldots\ldots k)(T = 0,1,2, \ldots). \quad (2.4)$$

In vector form, equation (2.3) becomes

$$n(T+1) = n(T)P + W(T+1) \quad (2.5)$$

When using equation (2.5) for forecasting, we start from the current transition counts of the CD_4 cells and then predict future transition counts for any time horizon, T, say provided that the transition probability matrix P is stationary over time or in other words is independent of time.

3.0 Application

To demonstrate the efficacy of our model, we apply it to a cohort study of 1094 HIV/AIDS positive patients with follow-up in their transition counts from January-December 2010.

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 patients CD_4 cell counts were classified into (5) states using the classification developed by Guiseppe Di Biase et al [4]. The classifications are summarized below.

State I: $CD_4 > 500$cells/mm^3
State II: $350 < CD_4 \leq 500$cells/mm^3
State III: $200 < CD_4 \leq 350$cells/mm^3
State IV: $CD_4 < 200$ cells/mm^3
State V: Death (Absorbing State).

4.0 Stationary Markov Chain Model

The transition matrices and transition probabilities of the 1094 patients of the cohort study for the twelve (12) observable months January-December 2010 are recorded as;

Table I: Represents the transition counts of the CD_4 cells of the 1094 patients in January 2010. Similar classifications were done for the month of February-December 2010 (See Appendix B).

JANUARY 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	95	99	48	18	12	272
$350 < CD_4 \leq 350$cells/mm^3	II	100	96	50	20	4	270
$200 < CD_4 \leq 350$cells/mm^3	III	70	50	95	55	7	277
$CD_4 \leq 200$cells/mm^3	IV	76	95	30	64	10	275
DEATH	V	0	0	0	0	0	0

Table II: Represents the total transition counts of the CD_4 cells of the cohort for the months of January-December 2010. By equation (2.3) we compute the transition probability matrix *P* for the stationary Markov chain model.

TOTAL TRANSITION COUNTS OF CD_4 CELLS

TOTAL TRANS. COUNTS		I	II	III	IV	V	TOTAL
$CD_4 > 500 cells/mm^3$	I	903	875	580	350	40	2754
$350 < CD_4 \leq 350 cells/mm^3$	II	857	936	584	376	41	2795
$200 < CD_4 \leq 350 cells/mm^3$	III	631	800	848	372	38	2689
$CD_4 \leq 200 cells/mm^3$	IV	631	880	785	371	45	2712
DEATH	V	0	0	0	0	0	0

Table III: Represents the transition probability matrix *P*.

$$P = \begin{array}{c} \\ I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{ccccc} I & II & III & IV & V \\ 0.328 & 0.318 & 0.211 & 0.127 & 0.016 \\ 0.307 & 0.335 & 0.209 & 0.135 & 0.014 \\ 0.235 & 0.298 & 0.315 & 0.138 & 0.014 \\ 0.233 & 0.324 & 0.289 & 0.138 & 0.017 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

The stationary Markov chain model is associated with the power 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 the probabilities are stationary over time. By equation (2.3), we compute the transition probability matrix *P* from the transition counts of the CD_4 cells from January-June 2010 and use it to test the efficacy of the model.

Table IV: Represents the transition probability matrix *P* from the months of January-June 2010.

$$P = \begin{array}{c} \\ I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{ccccc} I & II & III & IV & V \\ 0.328 & 0.318 & 0.211 & 0.127 & 0.016 \\ 0.307 & 0.335 & 0.209 & 0.135 & 0.014 \\ 0.235 & 0.298 & 0.315 & 0.138 & 0.014 \\ 0.233 & 0.324 & 0.289 & 0.138 & 0.017 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

By equation (2.5) we make predictions of the CD_4 cell counts for the month of July-December, 2010.

Table V: PREDICTION OF THE CD_4 CELL COUNTS FOR JULY-DEC. 2010

JULY 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 237 \\ 257 \\ 234 \\ 252 \\ 0 \end{array} \right)$$

AUGUST 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 235 \\ 236 \\ 234 \\ 234 \\ 0 \end{array} \right)$$

SEPTEMBER 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 223 \\ 224 \\ 222 \\ 222 \\ 0 \end{array} \right)$$

OCTOBER 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 217 \\ 218 \\ 216 \\ 216 \\ 0 \end{array} \right)$$

NOVEMBER 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 219 \\ 220 \\ 220 \\ 218 \\ 0 \end{array} \right)$$

DECEMBER 2010

$$\begin{array}{c} I \\ II \\ III \\ IV \\ V \end{array} \left(\begin{array}{c} 204 \\ 205 \\ 205 \\ 204 \\ 0 \end{array} \right)$$

These predicted values represents the actual transition counts of the CD_4 cells in each state in these months July-December 2010.

Table VI: Represents the comparison of the predicted CD_4 counts and the actual counts of CD_4 cells.

COMPARISON OF THE ACTUAL CD_4 CELL COUNTS AND THE PREDICTED CD_4 CELL COUNTS

		JULY	AUG	SEPT	OCT	NOV	DEC
STATE I	ACTUAL	258	231	232	256	242	224
	PREDICTED ABS	237	235	223	217	219	204
	ERROR	21	4	9	39	23	10
		JULY	AUG	SEPT	OCT	NOV	DEC
STATE II	ACTUAL	259	256	251	252	223	223
	PREDICTED ABS	257	236	224	218	220	205
	ERROR	2	20	27	34	3	18
		JULY	AUG	SEPT	OCT	NOV	DEC
STATE III	ACTUAL	228	228	237	231	243	231
	PREDICTED ABS	234	234	222	216	220	205
	ERROR	6	6	15	15	23	26
		JULY	AUG	SEPT	OCT	NOV	DEC
STATE IV	ACTUAL	224	249	238	238	237	242
	PREDICTED ABS	252	234	222	216	218	204
	ERROR	28	15	16	22	19	38
		JULY	AUG	SEPT	OCT	NOV	DEC
STATE V	ACTUAL	0	0	0	0	0	0
	PREDICTED ABS	0	0	0	0	0	0
	ERROR	0	0	0	0	0	0

MEAN ABSOLUTE ERROR (MAE) = 15.63

Table VII: Represents the predicted CD_4 cells for the month of January-June 2011.

PREDICTION OF THE TRANSITION COUNTS OF CD_4 CELLS

OVER THE MONTHS OF JANUARY-JUNE 2011

STATE	JAN	FEB	MARCH	APRIL	MAY	JUNE
I	195	268	258	250	239	229
II	195	269	259	251	240	230
III	195	270	259	251	240	230
IV	195	269	258	251	239	229
V	0	0	0	0	0	0

5.0 Conclusion

The stationary Markov chain model is applied in this study to capture AIDS dynamic progression of a patient. It is a multi-state model that considers the randomness of the time that the patient spends in a given state of the disease based on the epidemiological state of the CD_4 cell counts. The model allowed for the prediction of future CD_4 cell counts with a mean absolute error (MAE) of 15.43. As a predictive multi-state model, clinician and other health workers that are often confronted with the problem of future forecast of CD_4 cell counts may find the model useful in predicting future Biomarkers.

APPENDIX A
NOTATIONS

The notations used are defined as follows.

1. K = the number of states in the system
2. T = Calendar times in months, T = 0, 1, 2…
3. n_{ij} = the number of patients in state i in month T who transited into state j in month T + 1.
4. n_i = number of patients in state i at the beginning of the month T.
5. $n_{i,k+1}(T)$ = number of patients who died from state i at month T.
6. $P_{ij} = \frac{n_{ij}(T)}{n_i(T)}$ is the probability that a patient in state i transits to state j at the end of the month T.

7. $N(T) = \sum_{i=1}^{K} n_i(T) =$ Total number of patients at the end beginning of the period.
8. $W_i(T) = \frac{n_{i,K+1}(T)}{n_i(T)} =$ death rate of patients in state i at the month T.
9. $n_{0,j}(T+1) =$ new entrants into states j at the beginning of the month T.

APPENDIX B:

FEBRUARY 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	110	90	50	15	2	267
$350 < CD_4 \leq 500$cells/mm^3	II	80	120	55	15	5	275
$200 < CD_4 \leq 350$cells/mm^3	III	65	90	70	30	80	263
$CD_4 \leq 200$cells/mm^3	IV	60	84	85	20	1	250
DEATH	V	0	0	0	0	0	0

MARCH 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	95	86	50	20	9	260
$350 < CD_4 \leq 500$cells/mm^3	II	75	90	50	37	8	260
$200 < CD_4 \leq 350$cells/mm^3	III	70	95	80	24	6	275
$CD_4 \leq 200$cells/mm^3	IV	55	85	80	24	6	250
DEATH	V	0	0	0	0	0	0

APRIL 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	60	80	70	50	13	273
$350 < CD_4 \leq 500$cells/mm^3	II	60	70	65	57	10	262
$200 < CD_4 \leq 350$cells/mm^3	III	50	75	80	21	2	228
$CD_4 \leq 200$cells/mm^3	IV	60	70	85	25	10	250
DEATH	V	0	0	0	0	0	0

MAY 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	80	90	50	17	1	238
$350 < CD_4 \leq 500$cells/mm^3	II	90	100	55	10	4	259
$200 < CD_4 \leq 350$cells/mm^3	III	60	70	90	23	2	245
$CD_4 \leq 200$cells/mm^3	IV	50	90	85	19	1	245
DEATH	V	0	0	0	0	0	0

JUNE 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	90	80	47	40	1	258
$350 < CD_4 \leq 500$cells/mm^3	II	70	80	60	47	2	259
$200 < CD_4 \leq 350$cells/mm^3	III	50	65	75	37	1	228
$CD_4 \leq 200$cells/mm^3	IV	50	70	60	43	1	224
DEATH	V	0	0	0	0	0	0

DECEMBER 2010		I	II	III	IV	V	TOTAL
$CD_4 > 500$cells/mm^3	I	83	75	50	35	1	244
$350 < CD_4 \leq 500$cells/mm^3	II	75	60	50	35	3	223
$200 < CD_4 \leq 350$cells/mm^3	III	56	65	78	30	2	231
$CD_4 \leq 200$cells/mm^3	IV	58	86	60	35	3	242
DEATH	V	0	0	0	0	0	0

References

[1] Anderson R.M., Medly G.F., May R.M., Johnson A.M. (1986). *A Preliminary study of the transition dynamics of the human immunodeficiency virus (HIV), the causative agents of AIDS.* I.M.A. Journal of Mathematics. Applied to medicine and Biology vol. 3., no. 4., pp 229-263.

[2] Gain J. (1963): *Formulae for projecting Enrolments and degree awards in Universities*. Journal of Statistical Society A. 126, pp 409 - 440.

[3] Bartholomew D.J. (1963). *Stochastic models for social process.* 2nd Ed., Wiley, London.

[4] Guiseppe Di Biase, Gughelimo D. Amico, Arturo Di. Girolomo: Jacque Jansee, Stefano Jacobilli, Nicola Tinari, Raimando Manca (2007). *A stochastic model for HIV/AIDS Dynamic evolution.* Journal of Mathematical problems in Engineering Vol. 1, Article 65636.

[5] Bruce A. Craig and Peter P. Sendi (2002). *Estimation of the transition matrix of the discrete time Markov chain.* Department of statistics, Purdue University West Lefeyette U.S.A. Journal of Health Economics, Vol. 5, pp 33-42.

[6] Cousine M.S., Shickle L.M., Bander J.A. (2002). *An introduction to predictive modeling of disease management risk stratification.* Journal of Disease management, Vol. 41, pp 157-165.

[7] Annah M. Jelfrey, Xiahua Xia (2003). *When to initiate HIV Therapy. A control theoretic Approach.* IEEE Transaction on Biomedical Engineering Vol. 50, No. 11, pp 22 – 45.

[8] Osisiogu U.A. (2004). *Application of Non Stationary Markov chain model to a Graded Manpower System in Oyo State*. Nigerian Journal of Science Vol. 38, pp 39-51.

[9] Osisiogu U.A. (2005). *Application of Probabilistic Method in Analysis of career prospect of staff of Oyo State civil service structure*. Journal of Information Communication and Computing Technologies, Maiden Edition, pp 81-87.

[10] Osisiogu U.A. (2006). *Predicting secondary school Enrolment by Non-stationary Markov chain model. A case study of secondary school enrolment in Ogun State.* Journal of information communication and computing technologies. Vol. 2, pp 90-94.

[11] Scott D. Grimshaw, William P. Alexander (2010). *Markov chain models for diseases, transition matrix estimation and forecasting*. Journal of Theoretical and Applied Information Technology, Vol. 2, pp. 301-339.

[12] Milaslave Hajek and Yashik Singh (2011): *Medical A1-HIV/AIDS: Treatment management system.* Proceeding Extended Abstract presented at the international conference, Beyond A1, 2011 Pilsen Czech Republic.