HIV -Forecasting Models of CD4 Count

Maheshappa K., Wali.K.S, Basavarajaiah.D.M

Department of Statistics, Gulbarga University, Karnataka, India

Abstract- A total of 497 HIV patients were selected from cohort retrospective and non concurrent method and classified in to "low group or medium group", HAART was started based on the CD4 count, (Threshold CD4 count <250 cells /microL and >250 Cells /microL).Associated risk factors were analyzed by the conditional logistic regression. The change in CD4 count was evaluated. The mean age of the HIV patients was 32.90 years and the majority (63%) were men. 484 patients were initiated ART at a CD4 <250 (IQR 49-125) and 14 patients initiated at CD4 >250 (IQR 250-375).Patients initiated at CD4 more than 250 were 12.50% less likely to die (aHR 0.31,95% CI- 0.19-0.49) and 20% less likely to be lost to follow up (aHR 0.59,95% CI- 0.41 - 0.84).Using subjective decision CD4 count as predictor variable (The variable CD4 count less than 250 at start ART and > 250 no start ART were coded ; Code 01 and code -0 respectively). The model was constructed by an iterative maximum likelihood procedure -2 log likelihood statistics was very small (107.49) with Cox Snell R^2 (0.013) and mean CD4 count was (458±36.10 µ/dl). It depicts that better the CD4 count at start of ART more the longevity of life.

Index Terms- HIV, HAART, CD4, MLE's, ART

I. BACK GROUND

The HIV –forecasting disease modeling is that it leads to clear statements of the assumptions about the biological and social mechanisms, which influence disease spread and dynamism. The model formulation process is more valuable for statistician, Epidemiologists, Mathematician's and modelers, because it forces them to be precise about the relevant aspects of diseases transmission, course of infectivity, recovery, treatment prognosis and renewal of susceptibility. .The Statisticians need to formulate the models clearly and precisely using different clinical and biological parameters, which has been very well-understood in connection to dynamics of HIV –diseases, such as increase in CD4 count after inception of HAART, spectrum of HIV TB co infection, HCV HBV co infection etc., The Complete statements of the assumptions have crucial role, so that the reasonableness of the model can be interpreted by the relevant conclusions. The matrix of the Forecasting model may help physician, policymakers, young researcher and innovators. Limited number of HIV-forecasting model's study has been documented in India .In this context the Present study aims to formulate the different forecasting models, assessing quantitative conjectures and identify the trend of CD4 count before initiation of HAART therapy.

II. MATERIALS AND METHODS

A total 497 HIV infected pre- ART and alive and on ART patients were recruited with written consent. Retrospective cohort data viz., baseline (start of ART treatment), first year , second year , third year ,fourth year and fifth years cohort data was recorded systematically. As per the National guidelines we have classified the CD4 count as low CD4 count and medium CD4 count (Threshold CD4 count <250 cells /micro L and >250 Cells /microL) .Collected data were compiled by using SPSS-16.50 version. Logistic regression forecasting model, spearman rank correlation, polynomial curve fitting model were used to draw the relevant conclusion.

III. MODEL FORMULATION

Retrospective Secondary data were collected from different government ART centres in Bangalore city. Logistic Regression model was used to find out the variation among CD4 count of PLHIV and HIV associated parameters. we have assumed that 'n' is independent Bernoulli random variables of CD4<250 micro/dL and CD4>250 micro/Dl with observed value $(Y_1,Y_2\ldots..Y_n)$,where 'i=1,2….n and $(Xi_1, Xi_2\ldots..Xi_P, Xi_{P+1}\ldots\ldots Xi_{P+q})_1$ be a (p+q) vector of explanatory variables and denoted $X=(X_1, X_2, X_3 \ldots X_n)_1$ Let ß $\Pi_i= \Pi (X_i) =Pr (Y_i=1/X_i)$ be the event of probability for each observation of i=1, 2…n.

$\pi = (\pi_1, \pi_2, \pi_3, \ldots. \pi_n)$, then the logistic regression model is $logit(\pi_i) = X\beta\emptyset$

$$Logit(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right)_{=} X_1\beta$$

$$\text{Logit } (\pi_i) = \text{Log } \left(\frac{\prod i}{1-\prod i}\right) = X1i\beta$$

Where $(\beta = \beta_1, \beta_2, \beta_3 \ldots \beta_{p+q})$ is un known parameter vector, the joint probability of the observed y_o is a product of 'n' Bernoulli i function

$$L(\beta) = \left\{\prod_{I=1}^{n} \pi_i 0_i (1-\pi_i) 1 - y_{0i}\right\}$$

$$L(\beta) = \left\{ \frac{Exp(y_0, X_\beta}{\prod_{i=1}^{n}(1 + \exp X_i, \beta)} \right\}$$

Present study hypothesis was tested by using unconditional likelihood method (for maximizing the likelihood function).Our fitted model is in the form of **Logit (P)= βo+$\beta_j X_j$;** the parameters β_o, β_1…… β_k is typically estimated using maximum likelihood theory , let' n' denote the sample size and let (X_{i1},X_{i2}….X_{ky}) for the i th observation (i=1,2,……n) , treating each Y_i as the outcomes of independent random variable with success of probability P_i .

$$\prod_{i=1}^{n} p_i{}^{yi}(1-\pi_i)(1-y_{0i})$$

$$\frac{exp\{\sum_i yi\ (\beta_0 + \sum_j \beta j\ X_{ij}\}}{\pi_i(1 + \exp \beta_0 + \sum_j \beta j\ X_{ij}}$$

The maximum likelihood estimates (MLE'^{s}), (β_1, β_2….. β_k) that maximum were used to compute MLE'^{S} by using Newton rapson iteration method .the convent of present model became

$$\text{Log}\left(\frac{P}{1-P}\right) = \beta_o + \beta_1 X_1 + \beta_2 X_2 \ldots\ldots + \beta_k X_k$$

IV. RESULTS AND DISCUSSION

Tab (1): Base line characteristics of PLHIV'S.

SL	Variable	Male Mean±SD	Female Mean±SD	Pool	P-Value
1	Age (Year)	**36.24 ±6.96 (295-59.35%)**	**32.26±7.18 (202-40.64%)**	**35.30±6.19 (497)**	**0.029***
2	Base line CD4	**128.29±82.62 (295)**	**116.36±56.46 (202)**	**123.44±71.98 (497)**	**0.492^{Ns}**
4.	CD4-at the end of six month	**380.36±135.20 (287)**	**302.01±152.28 (196)**	**348.56±142.13 (483)**	**0.036***
5.	CD4-at the end of one year cohort	**428.86±45.65 (285)**	**372.28±96.73 (195)**	**405.87±66.40 (480)**	**0.006****
6	CD4-at the end of Two year cohort	**465.69±39.40 (282)**	**379.36±65.39 (193)**	**430.61±49.96 (475)**	**0.021***
7	CD4-at the end of Three year cohort	**502±30.61 (278)**	**410.00±49.50 (191)**	**464.53±38.30 (469)**	**0.085***
8	CD4-at the end of fourth year cohort	**534.89±28.33 (275)**	**445.00±41.29 (187)**	**498.50±33.57 (462)**	**0.011***
9	CD4-at the end of fifth year cohort	**542.16±20.02 (272)**	**451.00±36.98 (185)**	**505.25±27.97 (457)**	**0.003***

Baseline characteristics. A total of 497 PLHIV were considered for the study, out of which male comprises 295 (26%) and Female was 202(47%).The average age of patient was 35.30±6.19years, CD4 count was recorded at follow up interval , the mean base line CD4 count was 123.44± 71.98 micro/Dl(P=0.492) ,CD4 count at six month, one year ,two years follow up ,three years ,fourth years and fifth years was 348.56±142.13(P=0.036) micro per Dl,405.87±66.40(P=0.006) micro per Dl,430.61±49.96(P=0.021) microperdL,464.53±38.30 (P=0.085) micro perdL, 498.50±33.57(P=0.011) micro per Dl and 505.27 ±27.97 micro per Dl (P=0.003) respectively presented in Tab(1). 51.0% PLHIV were taking first line HAART therapy and only 4 () patients while on second line HAART, mortality rate was found to be 40 (8.08%).The mean body weight was 63 kg (70.25kg for men and 62.38 kg for women), and the mean hemoglobin was 13.8 g/dL (14.2 g/dLfor men and 12.6 g/dL for women). The median CD4 count was 330 cells/micro dL and median TLC was 1,692 cells/L.

Logistic regression Model: As per the fitted model actual predicted observed odds ratio of the base line CD4 Count was (OR) 0.35; 95% with confidence interval (CI) 0.26–0.59; *P*<0.0028]. Later initiation of ART therapy OR reported was (OR 1.16; 95% CI 1.03–1.3; *P*<0.02) and not statistically significant (p>0.05).After one year completion of HAART therapy CD4 count is directly proportional to the increase of Immunogical goal. The result shown that, CD4 count was improved after one year completion of HAART therapy (mean 405.87±66.40 micro/Dl (P=0.006),at the end of the five year cohort, mean CD4 count was 505.25±27.97 micro/Dl (P=0.003) it was clearly shown that ,immunogical factor (increase of CD4 count) is inversely proportional to the decline of RNA plasma viral load.

Tab (2).Association relation between CD4 count and age of the patients.

Clinical parameters		B	S.E.	Wald	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
							Lower	Upper
1	**Age**	-.053	.021	6.438**	.011	.948	.910	.988
2	**Cd4-Baseline**	1.92	.002	.018ns	.892	1.000	.997	1.003
3	**Cd4-at the end of the five year co hort**	.002	.001	4.766**	.029	1.002	1.000	1.003
4	**Regimen**	.377	.296	1.626**	.202	1.458	.817	2.603
5	**WHO-Stage**	-.440	.190	5.358**	.021	.644	.444	.935
6	**Blood group**	1.280	.996	1.653**	.199	3.597	0.652	0.889

2 Log likelihood, Cox & Snell R Square-0.076, Nagelkerke R Square-0.103

The associated influenced parameters for the increase of CD4 count was presented in Tab (2). As per the fitted model age β(-0.053) SE (0.021) CI 95 % 0.910-0.988 , CD4 count at base line (Inception of HAART) β(1.92) SE (0.002) CI 95 % 0.997-1.003, type of regimen (NRTI+2NNRTI, NNRTI+2NRTI) β(0.377) SE (0.296) CI 95 % 0.817-2.603 WHO clinical stage β(-0.440) SE (0.190) CI 95 % 0.44-0.935 and blood group β(1.280) SE (0.996) CI 95 % 0.652-0.889. The age, regimen, WHO clinical stage and Blood group were statistically significant (p<0.05) and more induced factors for the increase of CD4 count up to 192 micro per Dl

Tab (3).Correlation matrix of influenced parameters of HAART therapy.

Clinical parameters		Blood group	Age	Cd4-baseline	Cd4-present	Regimen	WHO-stage
01	Blood group	**1.000**					
02	Age	-.753	**1.000**				
03	Cd4-Baseline	-.088	-.062	**1.000**			
04	Cd4at fifth year cohort	**0.62****	**0.70****	-.266	**1.000**		
05	Regimen	-.386	-.047	-.240	**.742***	**1.000**	
06	WHO-Stage	-.429	**.082***	**.490***	**0.541***	.000	**1.000**

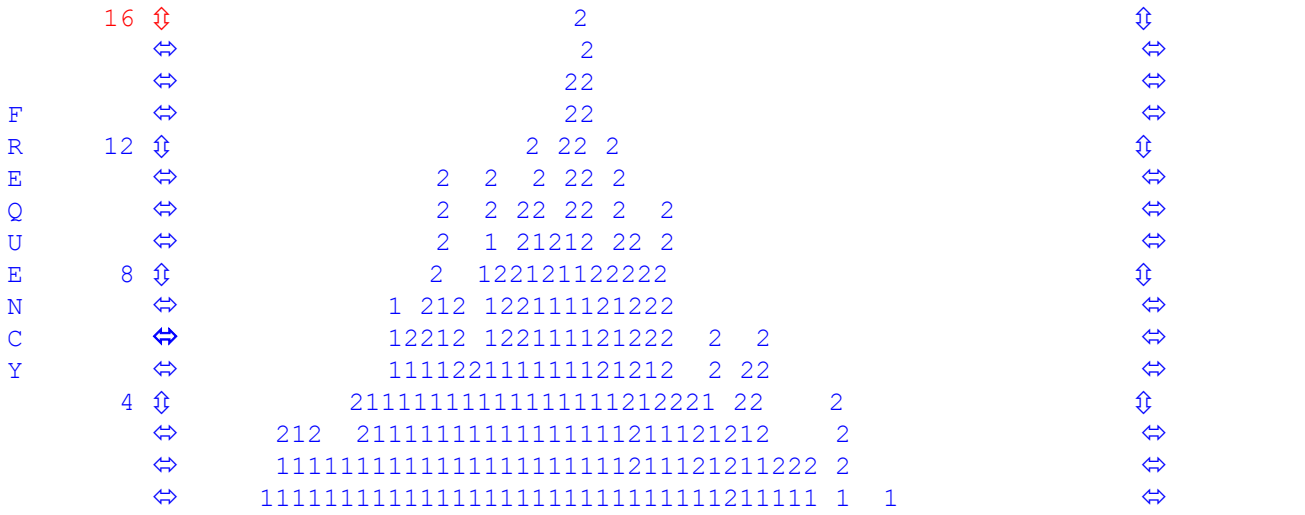
*, **Significant @0.05**

Correlation matrix of HAART therapy and concurrent parameters were analyzed by using spear man rank correlation method presented in Tab (3). Age(r= -0.753) ,CD4 count at inception of HAART(r= -0.088) , regimen (r= -0.386) and WHO clinical stage (r= -0.429) were negatively correlated with patient blood group.CD4 count(r= 0.70) and WHO Stage(Stage I,II,III and IV) (r= 0.082) were positively associated with age of the patients , regimen and advanced WHO clinical stage(r= 0.490) were positively correlated with CD4 count at the end of the five year cohort (r= 0.742).

V. CURVE FITTING

Extent and determinants of long-term increase of CD4 cell after initiation of antiretroviral therapy (ART) were compiled by using curve fitting model,. The mean base line CD4 was 128±82.62 cells/mm^3 (95% confidence interval [CI 95%105-136) in males and females was 116.36±56.46 cells/mm^3 (95% CI =96- 102.50) respectively .Independent predictors of long-term CD4 cell reconstitution variable were analyzed by linear and poly nominal Curve fitting method .Fifth and fourth degree poly nominal equation was fitted with respect to the gender .

Fig (1) : Observed Groups and Predicted Probabilities



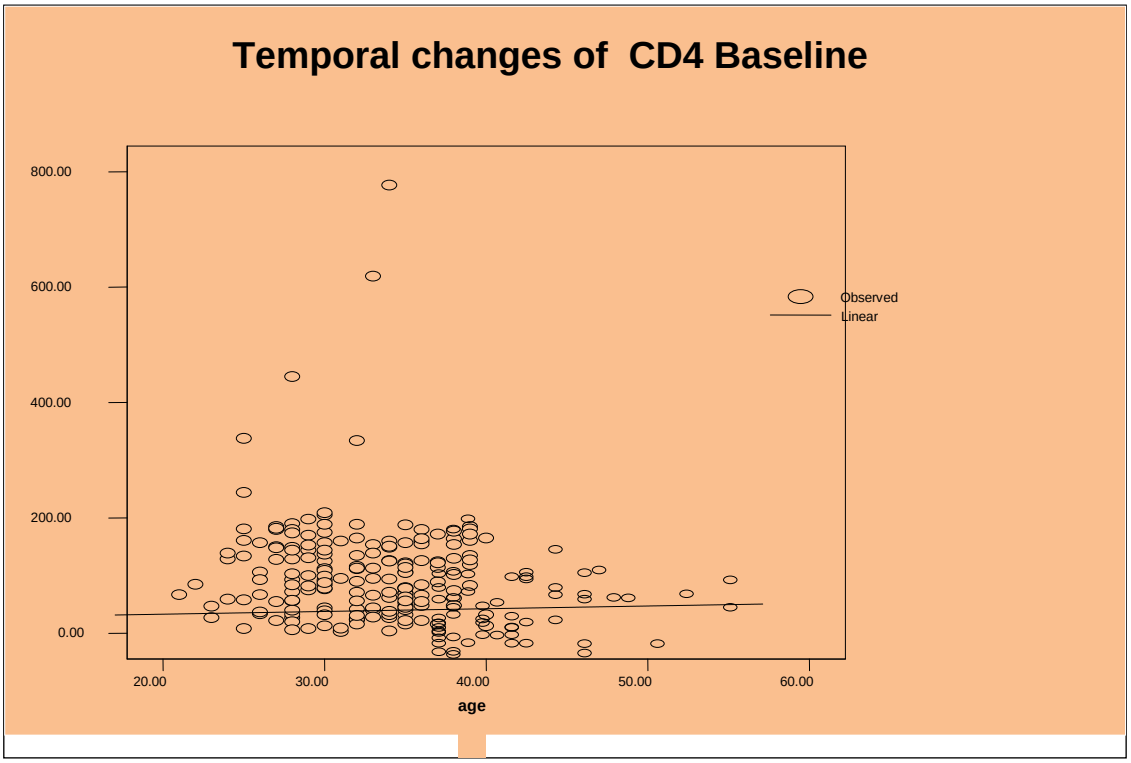


Fig (1): Curve Fitting-Linear trend (CD4 Count at base line).

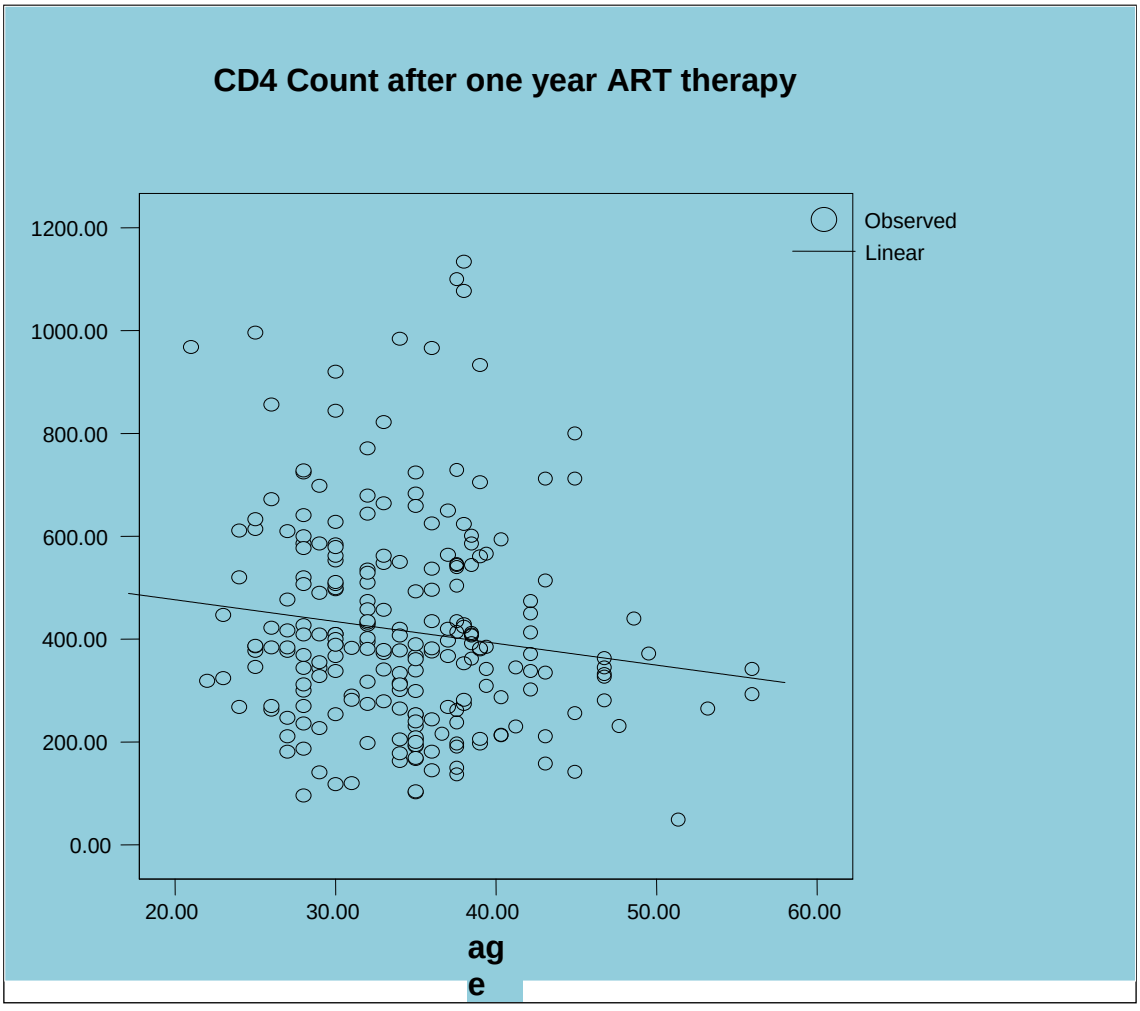


Fig (2): Curve Fitting-Linear trend (CD4 Count at the end of the cohort)

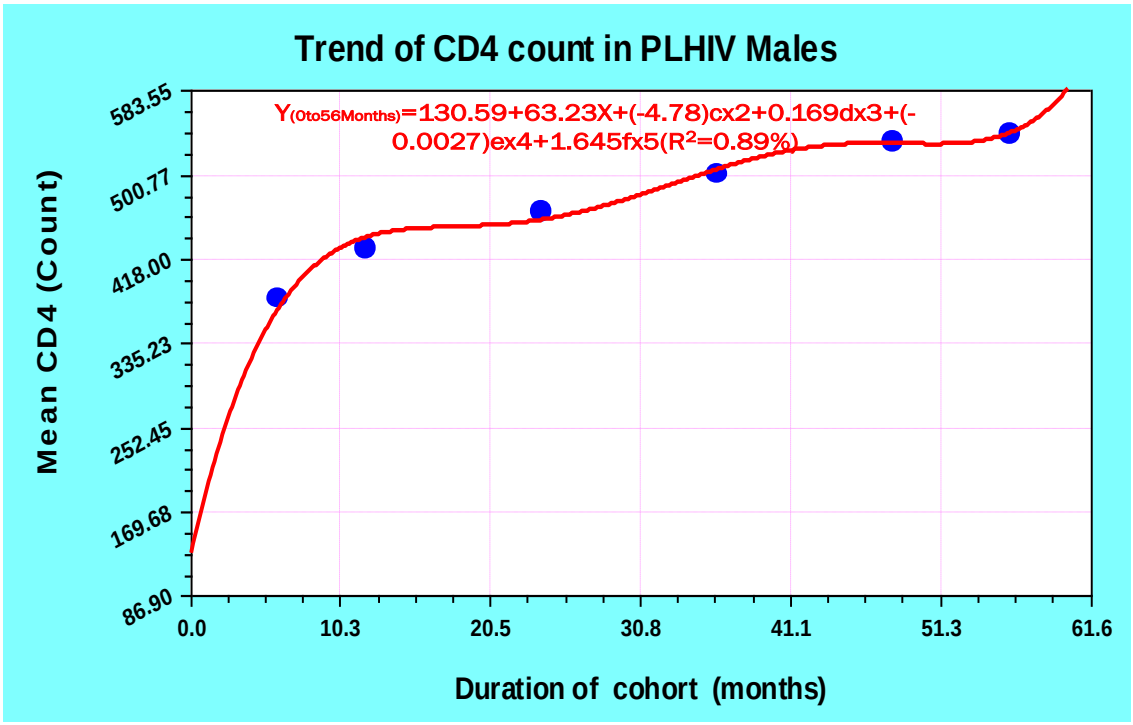


Fig (3) Fifth degree polynomial trend of CD4-Countin PLHIV Males (N=295)

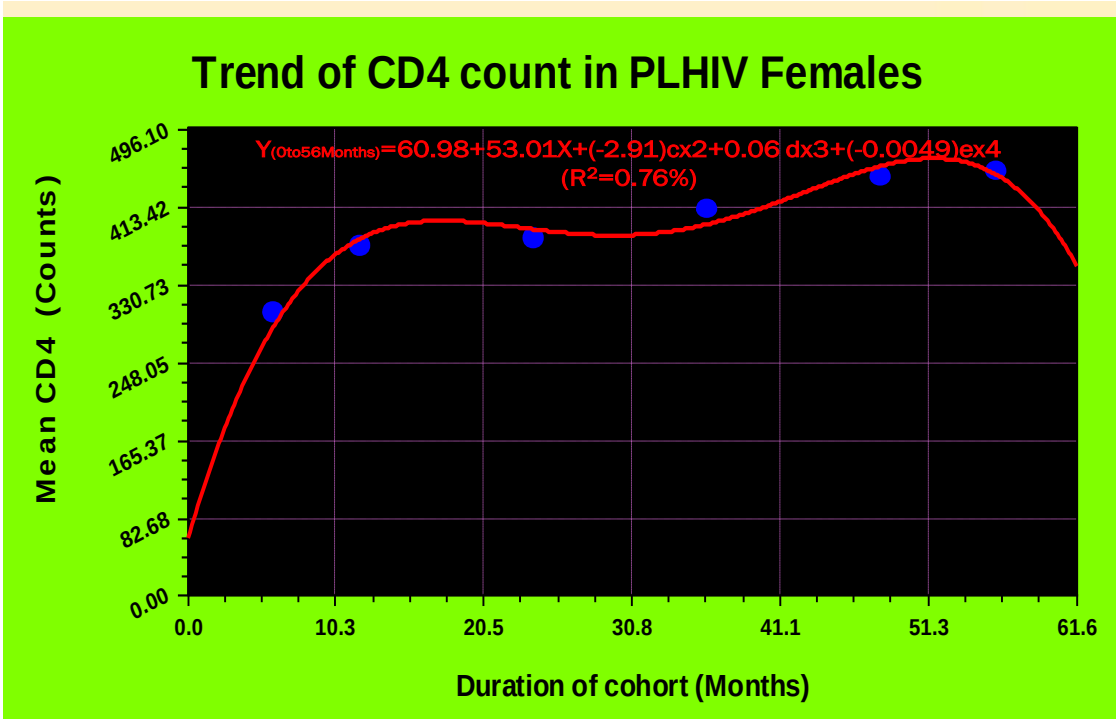


Fig (4) Fifth degree polynomial trend of CD4-Countin PLHIV Females (N=202)

The logit model reflects the 2×2 table

The Present study was demonstrated by using odds ratio model, the model expressed the ratio of the number of Patients who acquired CD4 count (<250 to >250 micro/Dl) with respect to the age of the patients (25-35 years). The formulation of the model was expressed by 2 × 2 contingency table:

$$\text{OR}=\left(\frac{AxD}{BxC}\right)$$

Tab(4): Association between age and base line CD4 count

CD4 Count	25-35 years	>36 years
CD4 Count <250 micro/dL	188 (A)	295 (B)
CD4 Count >250 micro/dL	09 (C)	05 (D)

$$\text{OR}=\left(\frac{188(A)x05(D)}{295(B)x09(C)}\right)=0.35$$

The model determines that, 65% of patients above the age group of 36 years have base line CD4 count <250µ/dl. The standard dispersion of CD4 count and age group of the patients was calculated. (Collett1991),

$$\textbf{SE ln (OR)} = \sqrt{\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)}$$

$$\textbf{SE ln (OR)} = \sqrt{\left(\frac{1}{188}+\frac{1}{295}+\frac{1}{9}+\frac{1}{5}\right)} \textbf{=0.564}$$

95%CI for the ln (OR) = −1.032 ± 1.96 × .564 = (−1.474,−.590)
95%CI for OR = (e−1.474, e−.590) = (.229, .554).

VI. DISCUSSION

As per the model formulation, the odds prediction equation was $ODDS = e^{a+bX}$. The fitted model R^2 was 80.26%($ODDS = e^{-.053+1.28(0)} = e^{-.931} = 0.45$. $ODDS = e^{-.053+1.28(1)} = e^{.37} = 1.458$). over a period of time the male and female cd4 count was found to be 45µ/dl and 14.58 µ/dl, which were likely to improve the CD4 count after initiation of HAART therapy. Due to adverse (ADV) and serious adverse drug reaction(SDV) the demonstrated model was $\hat{Y} = \frac{ODDS}{1+ODDS} = \frac{0.450}{1.458} = 0.30$. Based on this fitted model 30.80% and 59.0% of women and men enrolled for ever started HAART. Base line CD4 count was not statistically significant ($p<0.05$) and it was negatively correlated (r=-0.62),(-0.88) with clinically associated parameters viz., age of the person , WHO-clinical stage and HAART regimen. As per the fitted model, we have explained that , whose CD4 count was less at the time inception of HAART generally suffered from diarrhea ,fatigue, high fever, OI's and dermatological complications . After one year completion of HAART with effective counseling, palliative care and comprehensive support, theCD4 count was positively correlated with clinically associated parameters (r=+ve 0.70) presented in Tab (3).The variance of the estimated model was calculated by using MLE's. The main implications of our fitted model were designed to ensure to maintain maximum efficiency, accuracy and relevancy. However, our fitted model clearly demonstrated that, MLE's of the model-2 log likelihood was found to be 107.49 with very less expression of Coxsnell R^2. As per the above study, the results shown that the fitted model is considered as user-friendly due to relative prediction, efficiency, sufficiency and accuracy.

VII. CONCLUSION

The fitted model is very useful for clinical decisions, and also it can be used to predict the future trend of different HIV – Biomarkers. The method of ODD ratio model would provide good estimates of slopes and their standard errors.

ACKNOWLEDGMENTS

The authors acknowledge National AIDS Control organization, New Delhi, Project Director & Consultant (C.S.T) Karnataka AIDS Prevention society Bangaluru , Superintendent of Bowring &lady Curzon hospitals ,Department of Medicine ,B&LCH and Bangaluru Medical college & research Institute for giving permission to use HIV data to general analysis and given constant encouragement given to us for doing this kind of research.

REFERENCES

[1] Becton et al., CD4 cell count estimation were done by falcs count 2003; BMC Gastroenterology (5); PP 236-242.

[2] Vajpayee N,Kanswal S. Seth P, Wig N.Spectrum of opportunistic infection and profile of cd4 counts among AIDS patients in north India .*Infection* .2003;PP31:336-340.

[3] Framm S.R, Soave R. Agents of diarrhea. *Med Clin North Am*. 1997:427-447. doi: 10.1016/S0025-7125(05)70525-3.

[4] Joshi M, Chowdhry A.S Dalar PJ, Maniar JK.Prevelence of intestinal parasitic pathogens in HIV –seropositive individuals in northern India .*Natl Medj India* .2002:15:72-74.

[5] Lanjewar DN, Rodrigues C, Saple DG, Hira SK, Dupont. Crypto sodium, isospora and strongyloides in AIDS .*Natl Med J India*.1996; 9:17-19

[6] Mohan Das ,Sehgal R, Sud A, Malla N. Prevalence of intestinal parasitic pathogens in HIV –seropositive individuals in northern India .*Jpn J Infect Dis*. 2002;55:83-84.

[7] Curry A, Turner AJ, Lucas S. Opportunistic protozoan infections in HIV: Review highlighting diagnostic and therapeutic agents .*J Clin Pathol.1991:44:182-193.doi:10.1136/jcp.44.3.182*.

[8] Government of India. Epidemiology of HIV/AIDS.p 1-11 In B.B .Rewari (ed). Specialist *Training and reference Module*. National AIDS control Organization, New Delhi; 1999.

[9] Diagnostic Procedure for stool specimens. Http://www.dpd.cdc.gov

[10] Sengupta D, Lal S, Shrinivas Opportunistic Infection in AIDS. *Natl Med J India*.

[11] Guadalupe M, Sankaran S, George MD, Reay E, Verhoeven D, Shacklett BL, FlammJ, Wegelin J,Prindiville T, Dandekar S. Viral Suppression and Immune Restoration in the Gastrointestinal Mucosa of HIV Type 1-Infected Patients Initiating Therapy during Primary or Chronic Infection. *Journal of virology* .2006;80:8236-8247.doi:10.1128/JVI.00120-06.

[12] Yeni, HammerSM, Carpenter CC, Cooper DA, Fiscal MA,Gatell JM,et al.,.Antiretroviral treatment for adult HIV infection in 2002:updated recommendations of the International AIDS Society –USA Panel. *JAMA*2002; 288:222-235.

[13] Boerma JT, Holt E, Black R: Measurement of biomarkers in surveys in developing countries: Opportunities and problems. *Popul Dev Rev* 2001, 27(2):303-314.

[14] Garcia-Calleja JM, Marum LH, Carcamo CP, Kaetano L, Muttunga J, Way A: Lessons learned in the conduct, validation, and interpretation of national population based HIV surveys.

[15] UNAIDS/WHO: *Guidelines for measuring HIV prevalence in population-based surveys*. Geneva: UNAIDS/WHO Working Group on Global HIV/AIDS and STI Surveillance; 2005

[16] Ghys PD, Walker N, McFarland W, Miller R, Garnett GP: Improved data, methods and tools for the 2007 HIV and AIDS estimates and projections

AUTHORS

First Author – Maheshappa .K., PhD, Scholar, Department of Statistics, Gulbarga University, Gulberga
Second Author – Dr. Wali K.S. Chairmen and Professor and Head, Department of Statistics, Gulbarga University, Gulberga
Third Author – Mr.Basavarajaiah D.M. Biostatistician, Centre of Excellence in HIV care, Bowring and Lady Curzon hospitals, Shivajinagar, Bangalore-560001