WORLD JOURNAL OF PHARMACEUTICAL RESEARCH
SJIF Impact Factor 6.805
Volume 5, Issue 12, 975-985. Research Article ISSN 2277– 7105

MODELING OF HIV/AIDS SURVIVAL AND STAGING USING SURROGATE MARKERS

***P. Venkatesan**

Department of Community Medicine Sri Ramachandra University Porur, Chennai-600 116.

Article Received on
11 Oct. 2016,

Revised on 01 Nov. 2016,
Accepted on 21 Nov. 2016

DOI: 10.20959/wjpr201612-7480

***Corresponding Author**
P. Venkatesan
Department of Community Medicine Sri Ramachandra University Porur, Chennai-600 116.

ABSTRACT

For staging and survival, guiding treatment decisions and evaluating effectiveness of therapy of HIV/AIDS, CD4 testing is used as the gold standard. It is expensive to use in resource settings and there is a need for a less expensive surrogate marker to replace it. Several markers were studies in the past and Total Lymphocyte counts (TLC) was one the most studied worldwide. Since TLC varies according to ethnic groups, there is a need to assess the suitability for a particular population. This study sought to assess the suitability of TLC as a surrogate marker for CD4 count in Indian population. The data for the study was taken from 395 patients who were referred to a HIV clinical trial and tested for both CD4 count and TLC. The sensitivity, specificity and predictive values of various cut-off points of TLC were calculated. The decision theory based analysis was also carried out for various CD4 counts and their corresponding TLCs. The sensitivity and specificity of TLC in classifying HIV and AIDS based on CD4 as gold standard were about 50% and 95% respectively. Hence we conclude the TLC seems to be not an adequate marker for HIV/AIDS classification in Indian population as the sensitivity was very low.

KEYWORDS: TLC, CD4 count, HIV, AIDS, biomarker, Sensitivity, Specificity.

INTRODUCTION

The virologic load and immunologic surrogate markers such as plasma HIV RNA copies (viral load) and CD4+ cell counts play important roles in survival, staging and evaluating therapies in AIDS clinical research. The CD4+ cell counts serve as a primary surrogate marker. The HIV infection causes a gradual loss of CD4 T-cell counts and this is used to monitor HIV disease progression throughout the world.[1-3] It has been established that the normal rages of CD4+ T-cell counts vary among different ethnic populations.[4-5] The CDC

(1993) classification of HIV combining 3 categories of CD4 counts with 3 categories of symptoms is used to guide clinical and therapeutic actions in the management of HIV infected adolescences and adults.[6] The CD4 T-cell classification of CDC is based on 7 data sources which exhibited lot of variations. Hence it is necessary to have revised classification rules for different populations based on their CD4 T-cell distributions.

The CD4 count estimation is a costly method and serial CD4 count estimation for monitoring progression and ART evaluation is not practicable in limited resource settings. Even though the HIV RNA level (viral load) was shown to be more predictive than the CD4 counts, again the cost considerations make it not a viable alternative. More over the viral load and CD4 counts are negatively correlated during ART. It has been suggested that combination of both the markers may be a better and appropriate. If a new simple and cost effective method is developed as a surrogate marker for HIV disease progression, it has enormous impact on the treatment and management of HIV/AIDS patients.

The identification and improved use of intermediate markers of disease progression are therefore very important for early clinical decision making. One such surrogate marker considered worldwide is the Total Lymphocyte Counts (TLC).[7-8] It has been reported that it has poor specificity and/or low sensitivity. But the results were based on small samples and centered on specific databases. Very little work has been carried out in the direction of construction of probability models to optimize the sensitivity and specificity and prediction of CD4 from TLC. This work was an attempt towards that direction.

AIM

The main aim of this paper is to evolve a prediction model that predicts CD4 count from TLC with minimum error and to formulate a decision rule for HIV staging based on TLC and to evolve cut off points for Indian population.

METHODOLOGY

HIV preferred targets are T helper cells that has a docking molecule called “Cluster Designation 4” (CD4) on their surfaces. Cells with this molecule are known are known as CD4-positive (CD4+) cells. Destruction of CD4+ lymphocyte levels appears to be the best indicator for developing opportunistic infections.[9] The rate of decline of CD4 counts in HIV patients has been determined at different levels and different ranges. The form of decline through time has been characterized non-linear curves.

A probability model was constructed to discriminate HIV patients at different cutoff points that optimize sensitivity and specificity (A weighted combination of both). A weighted non-parametric approach was also used to formulate an optimal model that predicts CD4 from TLC with minimum error. The likelihood ratio method, receiver operating characteristic curve and precision-recall curve methodology were applied to optimize the classification error.

The CD4 counts and TLC are in continuous measurement scales. Let X denote the random variable denoting the continuous measurements. The k be threshold value of X used to classify patients such that if X<k, the patients are classified as AIDS and X ≥k, non-AIDS. If the true states of the patients are known, then we have the usual two way table at k.

Table 1: Validation Measures for a Marker against a gold standard

Marker	Gold Standard		
	Diseased	Non-Diseased	Total
X<k	A	B	A+B
X≥k	C	D	C+D
Total	A+C	B+D	N

The accuracy of such binary diagnostic test is commonly assessed by sensitivity and specificity. The sensitivity is the probability that the test correctly classifies a disease subject as diseased and specificity is the probability that the test correctly classifies non-diseased individual as diseased. If D denotes the true disease status, define D=1 as diseased and D=0 as non-diseased, then from Table 1.

Sensitivity = P(X+/D=1) = A/(A+C);

Specificity = P(X-/D=0) = B/(B+D)

False positive rate= P (D=0/X+) =B/(A+B);

False negative rate= P (D=1/X-) =C/(C+D) ---------- (1)

The likelihood ratios for diseased and non-diseased for any marker with cutoff point can be calculated as given in Table 2.

Table 2: Likelihood Ratio based on gold standard.

	Diseased	Non-Diseased	

Marker	No	Likelihood	No	Likelihood	Likelihood Ratio
<k	A	A/(A+C)	B	B/(B+D)	A(B+D)/(B(A+C)
≥k	C	C/(A+C)	D	D/(B+D)	C(B+D)/(D(A+C)
	A+C		B+D		

Based on decision theory approach, the expected gain in the CD4 counts can be calculated using the equation(2).

$$g(k) = \Sigma_i \Sigma_j w_{ij} p_{ij} \qquad i, j=1,2 \qquad \text{---------- (2)}$$

where p_{ij} are the weighted probabilities as given in Table 3. The optimum k can be obtained by evaluating g(k) for various values of k

Table 3: Decision theory weights

Marker	Gold standard		
	Diseased	Non-diseased	Total
<k	p11 (w11)	p12(w12)	p1.
>=k	p21(-w21)	p22(-w22)	p2.
Total	p.1	p.2	1

Model:_ When evaluating a continuous scale test, both sensitivity and specificity depends on the test threshold and varies according variation of k. By considering all possible values of the threshold for X, a receiver operating characteristic (ROC) curve can be constructed as a plot of sensitivity against (1-specificity). If $F_{d=(0,1)}$ is the distribution function of the continuous scale X, for a patient with D=d (d=0, 1), the ROC curve of X can be written as

$$ROC(p) = 1-F_1(F_0(1-p)) \qquad \text{---------- (3)}$$

Where the false positive rate corresponding to the cutoff point k is in the domain of the distribution function F_0. A vertical confidence interval for true positive rate(TPR)for a specified false positive rate(FPR) or a horizontal confidence interval for FPR for a specified TPR can be constructed often first in an unrestricted space and then transformed back to ROC space.[10] The transformations are carried out to improve large scale approximations. The non-parametric standard errors can be directly obtained or Greenhouse and Mantel type confidence rectangles or Kolmogorov –Smirnov type fixed width non-parametric bands can also be developed. The ROC gives equal weights to the full range of TPR and FPR where as only limited ranges may be of practical interest in AIDS studies. The partial area under the ROC between two fixed *a priori* values of specificities can be calculated.[11-13] The empirical ROC curve is always unsmooth and has a jagged form. Since the true ROC is a smooth function, the smoothing of non-parametric ROC curves can be derived from estimates of

density or distribution of two test distributions. The degree of smoothness is determined by the choice of kernel and bandwidth. A semi-parametric approach can be considered by assuming monotone transformation of X that simultaneously makes both distributions normal.

The Likelihood Ratio (LR) method can be used to generate an index which contrasts the proportions of patients with or without the target disorder given the level of a test. The likelihood ratio expresses the odds that a given diagnostic test result would be expected in a patient with as opposed to one without the disorder. The LR method has three important properties which when combined form a very powerful diagnostic strategy. First, the ratios are unaffected by the changes in the prevalence of the disorder and are more stable than sensitivity and specificity. Second, the LR can be calculated for several levels of the sign, symptom or test results rather than working at two levels. Thirdly the LR can be used in a powerful way to shorten a list of diagnostic hypotheses using the following equation:

Pre-test Odds * LR = Post-test Odds ---------- (4)

The post-test odds can be converted to post-test probability using the formula:

Post-test Probability= Post-test-Odds/ (Post-test-Odds+1) ---------- (5)

Database

The database consists of 395 HIV patients with TLC and CD4 counts referred to a clinical trial at the National Institute for Research in Tuberculosis (formerly Tuberculosis Research Centre), Chennai.[14] The database is used to make empirical comparison between CD4 and TLC in classifying AIDS from HIV. Based on WHO criteria CD4<200 was taken as AIDS patients and CD4>200 as non-AIDS (HIV) patients. The aim is evaluate WHO criteria of TLC<1250 is a surrogate marker for CD4<200 and their practical utility in determining the optimal cutoff points based on CD4/TLC.

RESULTS

The mean (SD) of CD4 counts for different ranges of TLC are given in Table 4. From the table we see that the mean CD4 counts increase as the TLC increases. The mean CD4 was 117 for TLC less than 1250. Table 5 gives the mean (SD) of TLC for difference ranges of CD4 counts. It is seen that for CD4 <200, the mean TLC was about 2000 much higher than the WHO cut off value 1250. The variances of CD4 on different ranges of TLC and variances of TLC on different ranges of CD4 varied markedly.

Table 4: Mean (SD) CD4 counts for different ranges of TLC

	CD4 counts				
TLC Group	**N**	**Mean**	**S.D**	**Min.**	**Max**
<500	4	40.0	35.6	8	80
500-750	23	47.4	40.4	6	151
750-1000	14	105.0	80.4	16	315
1000-1250	55	117.0	83.7	12	341
1250-1500	27	170.7	99.0	39	351
1500-1750	54	240.1	133.7	34	600
1750-2000	27	309.1	190.8	36	684
2000- 2250	45	324.3	163.2	42	792
2250-2500	28	312.2	171.6	72	696
>=2500	118	519.4	328.2	60	1980
Total	395	303.0	265.1	6	1980

Table 5: Mean (SD) TLC count for different ranges of CD4.

CD4 Group	**N**	**Mean**	**S.D**	**Min.**	**Max**
<50	44	956.8	415.1	400	2100
50-100	47	1506.4	761.4	400	3300
100-150	40	1577.5	733.3	500	3600
150-200	41	2036.6	1073.7	700	5100
200-250	31	1832.3	593.5	1100	3500
250-300	33	2178.8	765.6	1100	4900
300-350	34	2361.8	1004.5	900	4900
350-400	18	2127.8	534.5	1300	3400
400-450	17	2235.3	525.5	1500	3500
450-500	17	2576.5	716.4	1500	3800
500-750	49	2698.0	760.3	1500	5200
>=750	24	3979.2	1601.9	2200	9600
Total	395	1085.2	1085	400	9600

The proportions of patients whose CD4 counts are less than200 for different ranges of TLC are given in Table 6. When TLC was less than 750, 100% of the patients had CD4 less than 200 and 92.9% for TLC range (750-1000) and 83.6% for TLC range (1000-1250).

Table 6: Percentage of patients CD4<200 among different ranges of TLC

TLC Group	**CD<200**
<500	100.0
500-750	100.0
750-1000	92.9
1000-1250	83.6
1250-1500	63.0
1500-1750	40.7
1750-2000	33.3
2000- 2250	24.4

2250-2500	32.1
>=2500	15.3
Total	43.5

The proportion of patients who's TLC was less than WHO cut off point of 1250 are given in Table 7. It is found that only 84% were TLC less than 1250 even in the extreme cases of CD4<50. For CD4 (150-200) it was only 24% indicating the wide variability in TLC.

Table 7: Percentage of patients TLC <1250 among different ranges of CD4

CD4 Group	TLC<1250
<50	84.1
50-100	51.1
100-150	37.5
150-200	24.4
200-250	12.9
250-300	9.1
300-350	8.8
350-400	0.0
400-450	0.0
450-500	0.0
500-750	0.0
>=750	0.0
Total	24.3

Even though there is an increasing trend in the mean values of CD4 for increase in TLC, the range varies widely which makes the prediction of the regression model poor. Using the WHO classification rule of TLC<1250 for CD4<200, Table 8 presents the classification accuracy of TLC, considering CD4 as gold standard. The TLC marker has high specificity, but low sensitivity indicating that it is not an adequate marker to replace CD4. The positive and negative predictive values were moderate.

Table 8: TLC vs CD4 counts agreement

	CD4 counts			
		<200	≥200	Total
TLC	<1250	86	10	96
	≥1250	86	213	299
	Total	172	223	395

Sensitivity=50.0%, Specificity=95.5%, PPV=89.6%, NPV= 71.2%.

A weighted regression model was also fitted for the data and equation is given by

CD4= 53.7 + 0.132 * TLC ------- (6)

The R value of 0.57, indicating that the weighted model fit is also not good. The predicted CD4 value for TLC 1250 was about 220 close 200, but the 95% confidence limits were (165, 272) which is very wide indicating the limitation of the model for predicting CD4 given TLC.

The likelihood ratios for TLC using CD as gold standard arealso given in Table 9. The likelihood ratio of 11.4 for TLC <1250 suggesting that the chance of CD4<200 is high. There is also high chance that CD will be less than 200 if TLC value is less than 1500.

Table 9: Likelihood ratios for TLC (CD4<200 Vs CD4≥200)

TLC	Likelihood Ratio
<1250	11.4
1250-	2.1
1500-	0.9
1750-	0.6
2000-	0.4
>=2250-	0.3

The loss function is the penalty paid when κ' is not the estimate of k. Bayes estimates for a given loss function is one that minimizes the expected loss given the loss posterior distribution of the parameter: There are three common loss function approaches namely: Squared loss function (Bayes estimate is the mean of the posterior distribution), Absolute loss function (Bayes estimate is the median of the posterior distribution) and Zero-one loss function (Bayes estimate is the mode of the posterior distribution). The loss function approach also gave similar results indicating TLC is a poor predictor of CD4.

DISCUSSION

Depletion of CD4 cell subset of TLC has been noted as the hallmark of HIV infection and CD4 count has been established as the gold standard for survival prediction, staging of HIV/AIDS and guiding treatment decisions. In areas where viral load and CD counts are not available WHO guidelines the use of TLC in conjunction with clinical data as a criteria for initiation of ART. In this study threshold analysis was done to find the ability TLC to predictCD4 counts at two levels namely CD4<200 and CD4≥200.

The study found that a TLC of <1200 cells/mm^3 had a sensitivity of 50% and specificity of 95.5% for a CD4 count<200 cells/mm^3 with a PPV of 89.6% and NPV of 71.2%. These

results indicate that the relationship between CD4 and TLC is only moderate. TLC at a cutoff of 1200 seems to be not a good surrogate marker for CD4 in this population. This finding is in consistent with others.[15-17] However there have been conflicting results that has been reported.[18-20] These differences could be due to ethnic, racial, epidemiological and socioeconomic factors.[21]

CONCLUSION

The findings of the study suggest that TLC is not an accurate tool that can serve as a surrogate marker in HIV patients. Further studies are needed to construct models adjusting for other confounding variable to increase the predictive values of TLC.

Limitation

The data base is part patients referred of major HIV study and the subjects were naïve for ART. Interpretation of the should be done with caution.

REFERENCES

1. Prentice, RL. Surrogate end points in clinical trials: Definition and operational criteria. Statistics in Medicine, 1989; 8: 431-40.
2. Fleming, TR, DeMets, DL. Surrogate end points in clinical trials: Are we being misled? Annals of Internal Medicine. 1996; 125: 1014-29.
3. DeGruttola, VG, Clax, P., DeMets, DL, Downing, GJ, Ellenberg, S, Friedman, L., Gail, MH, et al. Consideration in evaluation of surrogate endpoints in clinical trials: Summary of a NIH workshop. Controlled Clinical Trials. 2001; 22: 485-502.
4. Crowe, S., Turnbull, S., Oelrichs, R. and Dunne, A. Monitoring Human Immunodeficiency virus infection in resource-constrained countries. Clinical Infectious Diseases. 2003, 37: S25-35.
5. Kam, KM Leung, WL, Kwok, MY, Hung, MY, LeeSS, Mak WP. Lymphocyte subpopulation reference ranges for monitoring human immunodeficiency virus infected Chinese adults. Clinical Diagnostic Laboratory Immunology, 1996; 3: 326-3.
6. CDC. 1992 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescentsand adults. MMWR, 1993; 41(no. RR–17): 961–2.
7. Post FA, Wood R, and Maartens G. CD4 and total lymphocyte counts as predictors of HIV disease progression. Quartly Journal of Medicine, 1996; 89: 505-08.

8. Shaprio NI, karras, DJ, Leech SH and Heilpern KL. Absolute lymphocyte count as a predictor of CD4 count. Annals of Emergency Medicine. 1998; 32: 323-28.
9. Beck EJ, Kupek Ej, Gompels MM and Pinching AJ. Correlation between total and CD4 lymphocyte counts in HIV infection: not making the good an enemy of the not so perfect. International Journal of STD and AIDS. 1996; 7: 422-28.
10. Blatt SP, Lucey CR, Butzin CA Hendrix CW and Lucey DR. Total lymphocyte count as a predictor of absolute CD4+ count and CD+ count and Cd4+ percentage in HIV infected persons. 1993; 269: 622-26.
11. Fournier AM and Sosenko, JM. The relationship of total lymphocyte count to CD4 lymphocyte counts in patients infected with human immunodeficiency virus. American Journal of Medical Science, 1992; 304: 79-82.
12. Lang W, Perkins H, Anderson RE, Royce R, Jewell N, Winkelstein W Jr. Patterns of T-lymphocyte changes with human immunodeficiency virus infection: from seroconversion to the development of AIDS. Journal Acquired Immuno Deficiency Syndrome, 1989; 2: 63–69.
13. Masur H, Ognibene FP, Yarchoan R *et al.* CD4 counts as predictors of opportunistic pneumonias in human immunodeficiency virus(HIV) infection. Ann Intern Med, 1989; 111: 223–31.
14. Soumya Swaminathan, Prdeep A.Menon, Narendran Gopalan, Venkatesan Perumal. et al., Efficacy of a Six Month Vesus a 36-Month Regimen for Prevention of Tuberculosis in HIV-Infected Persons in India: A Randomized Clinical Trial. *PLOS ONE*. 2012; 7(e47400): 1-9.
15. Daka D, Loha E: Relationship between total lymphocyte count (TLC) and CD4 count among peoples living with HIV, Southern Ethiopia: a retrospective evaluation. AIDS Res Ther, 2008; 5: 26.
16. Akinola NO, Olasode O, Adediran IA, Onayemi O, Murainah A, Irinoye O, Elujoba AA, Durosinmi MA: The search for a predictor of CD4 cell count continues: total lymphocyte count is not a substitute for CD4 cell count in the management of HIV-infected individuals in a resource-limited setting. Clin Infect Dis., 2004; 39(4): 579–581.
17. Angelo AL, Angelo CD, Torres AJ, Ramos AM, Lima M, Netto EM, Brites C: Evaluating total lymphocyte counts as a substitute for CD4 counts in the follow up of AIDS patients. Brazilian Journal Infectious Disease. 2007; 11: 466–470.
18. Obirikorang C, Quaye L, Acheampong I. Total lymphocyte count as a surrogate marker for CD4 count in resource-limited settings. BMC Infectious Diseases, 2012; 12: 128.

19. Post FA, Wood R, Maartens G: CD4 and total lymphocyte counts aspredictors of HIV disease progression. QJM, 1996; 89: 505–508.
20. Akinola NO, Olasode O, Adediran IA, Onayemi O, Murainah A, Irinoye O, Elujoba AA, Durosinmi MA: The search for a predictor of CD4 cell count continues: total lymphocyte count is not a substitute for CD4 cell counting the management of HIV-infected individuals in a resource-limited setting. Clinical Infectious Disease, 2004; 39: 579–581.
21. Alavi SM, Ahmadi F, Farhadi M: Correlation between Total Lymphocyte Count, Hemoglobin, Hematocrit and CD4 Count in HIV/AIDS Patients. Acta Medica Iranica, 2009; 47.