Complete Blood Cell Count as a Surrogate CD4 Cell Marker for HIV Monitoring in Resource-Limited Settings

Ray Y. Chen, MD,† Andrew O. Westfall, MS,‡ J. Michael Hardin, PhD,‡ Cassandra Miller-Hardwick, BSN,† Jeffrey S. A. Stringer, MD,§ James L. Raper, DSN, CRNP, JD,* Sten H. Vermund, MD, PhD,*‖ Eduardo Gotuzzo, MD,¶ Jeroan Allison, MD,* and Michael S. Saag, MD**

Background: A total lymphocyte count (TLC) of 1200 cells/mL has been used as a surrogate for a CD4 count of 200 cells/μL in resource-limited settings with varying results. We developed a more effective method based on a decision tree algorithm to classify subjects.

Methods: A decision tree was used to develop models with the variables TLC, hemoglobin, platelet count, gender, body mass index, and antiretroviral treatment status of subjects from the University of Alabama at Birmingham (UAB) observational database. Models were validated on data from the Birmingham Veterans Affairs Medical Center (BVAMC) and Zambia, with primary decision trees also generated from these data.

Results: A total of 1189 patients from the UAB observational database were included. The UAB decision tree classified a CD4 count ≤200 cells/μL as better than a TLC cut-point of 1200 cells/mL, based on the area under the curve of the receiver-operator characteristic curve ($P < 0.0001$). When applied to data from the BVAMC and Zambia, the UAB-based decision tree performed better than the TLC cut-point of 1200 cells/mL (BVAMC: $P < 0.0001$; Zambia: $P = 0.0009$) but worse than a decision tree based on local data (BVAMC: $P \leq 0.0001$; Zambia: $P \leq 0.0001$).

Conclusion: A decision tree algorithm based on local data identifies low CD4 cell counts better than one developed from a different population or a TLC cut-point of 1200 cells/mL.

Key Words: AIDS, $CD4^+$ cell count, complete blood cell count, decision tree, HIV, total lymphocyte count

(*J Acquir Immune Defic Syndr* 2007;44:525–530)

As highly active antiretroviral therapy becomes more accessible and affordable in developing countries through pharmaceutic company price reductions, generic drugs, and programs such as the Global Fund to Fight AIDS, Tuberculosis, and Malaria and the President's Emergency Plan for AIDS Relief, the cost of monitoring HIV therapy may become more prohibitive than the cost of the medications themselves.[1,2] The World Health Organization (WHO) recommends using a total lymphocyte count (TLC) of 1200 cells/mL as a surrogate marker for a $CD4^+$ count of 200 cells/μL for treatment initiation when $CD4^+$ cell counts are unavailable.[3] Many studies have evaluated the use of TLC as a surrogate marker for $CD4^+$ cell count with mixed results.[4,5] Some studies have found a good correlation,[6–17] but others have not.[18–22] In addition to low lymphocyte count, anemia,[23] thrombocytopenia,[24] and body mass index (BMI)[25,26] have been associated with advanced HIV infection. In an effort to improve the TLC model, other studies have added such parameters.[27–31]

In this study, we hypothesized that a decision tree algorithm based on multiple components of a complete blood cell count (CBC) and other easily obtainable variables would be more effective in identifying patients with a $CD4^+$ count ≤200 cells/μL than a decision rule based solely on a TLC cut-point of 1200 cells/mL. Furthermore, we examined the transportability of decision tree models across populations.

METHODS

The University of Alabama at Birmingham (UAB) Outpatient HIV Clinic began collecting information on all patients in an ongoing observational database in January 1994. Trained medical records personnel use standardized procedures to collect clinical and treatment data from medical records daily. Laboratory data are downloaded from the

Received for publication June 29, 2006; accepted December 20, 2006.

From the *Department of Medicine, University of Alabama at Birmingham, Birmingham, AL; †Birmingham Veterans Affairs Medical Center, Birmingham, AL; ‡Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL; §Department of Obstetrics and Gynecology, University of Alabama at Birmingham, Birmingham, AL; ‖Department of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL; and ¶Department of Transmissible Diseases, Hospital Cayetano Heredia, Lima, Peru.

R. Y. Chen is currently affiliated with the Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD. C. Miller-Hardwick is currently affiliated with Extended Care and Rehabilitation Service Line, Minneapolis Veterans Affairs Medical Center, Minneapolis, MN. S. H. Vermund is currently affiliated with the Institute for Global Health, Vanderbilt University School of Medicine, Nashville, TN.

Supported in part by grants from Bristol-Myers Squibb Virology, the Elizabeth Glaser Pediatric AIDS Foundation (PG 51161-26), the Department of Veterans Administration/Birmingham Veterans Affairs Medical Center, and the University of Alabama at Birmingham Center for AIDS Research (P30-AI27767).

R. Y. Chen received a grant from Bristol-Myers Squibb Virology. M. S. Saag has served as a consultant to or on the speakers' bureaus of Abbott, Bristol-Myers Squibb, Gilead Sciences, GlaxoSmithKline, Hoffman-La Roche, Ortho Biotech, Pfizer, Shire, Tibotec, Trimeris, and Monogram Biosciences.

Reprints: Michael S. Saag, MD, University of Alabama at Birmingham HIV Outpatient Clinic, Community Care Building, Room 142, 908 20th Street, South Birmingham, AL 35294–2050 (e-mail: msaag@uab.edu).

hospital laboratory system directly into the database. Outside laboratory values are entered into the database manually. The UAB Institutional Review Board (IRB) has approved the database protocol. Patients were included in this study if they had a complete set of laboratory data ($CD4^+$ cell count, TLC, hemoglobin, and platelet count) drawn on the same day. For patients with more than 1 set of complete laboratory data, 1 set was randomly selected for this analysis.

SAS Enterprise Miner (version 5.2; SAS Institute, Cary, NC) software was used to develop a decision tree model to discriminate a $CD4^+$ cell count $\leq 200/\mu L$. The variables used to develop the decision tree were limited to a minimal number of easily obtainable variables to make the model specific yet still broadly applicable. The variables included in this analysis were TLC, hemoglobin, platelet count, gender, BMI, and any antiretroviral therapy in the previous 30 days (yes/no). In accordance with standard practices, the data were randomly split into training and validation data sets. The random samples were stratified on the outcome variable to ensure comparability with respect to classification error rates. Seventy percent of the data were allocated to the training sample and the remaining 30% to the validation sample. The decision tree algorithms used were the default Enterprise Miner algorithm, the classification and regression tree (CART) algorithm, and the χ^2 automatic interaction detector (CHAID). The CART[32] algorithm uses a recursive partitioning Gini reduction strategy, whereas the CHAID[33] uses χ^2 tests to determine the best cut-points to use in developing the decision tree.

To determine the best algorithm, receiver-operator characteristic (ROC) curves were generated from the decision trees by varying the cutoffs, 1-by-1, for predicting a CD4 count ≤ 200 cell/mL. The first point on the ROC curve occurs when none of the nodes predicts a CD4 count ≤ 200 cells/mL [point (0,0); sensitivity 0%, 1-specificity 0%]. Each subsequent point on the ROC curve is determined by sequentially moving 1 terminal node to the predict ≤ 200-cells/mL group, gradually increasing sensitivity and 1-specificity until all nodes predict a CD4 count ≤ 200 cells/mL [point (1,1); sensitivity 100%, 1-specificity 100%]. The best algorithm, based on the largest area under the curve (AUC) of the ROC curve, was used to develop the decision tree model for this study.

The decision tree developed with UAB data was then used to discriminate a $CD4^+$ count ≤ 200 cells/μL on data from the Birmingham Veterans Affairs Medical Center (BVAMC) and Lusaka, Zambia, after receiving appropriate IRB approvals. Separate decision trees were also developed using the BVAMC and Zambian data directly to determine if a "custom-designed" decision tree classified better than a tree developed using data from a different (UAB) population. Because of the smaller sample size of the BVAMC and Zambian data sets, each data set was used in its entirety to develop the decision tree model rather than splitting into training and validation samples.

When 2 or more ROC curves are constructed based on the same individuals, statistical analysis on differences between curves must account for the correlated nature of the data. A nonparametric approach was used for the analysis of areas under correlated ROC curves using the theory of generalized U-statistics to generate an estimated covariance matrix.[34] For each site, a nonparametric comparison of the correlated AUCs was performed using the ROC macro available on the SAS Institute Web site.[35] The ROC macro performs statistical tests of the equality of all areas and pairwise comparisons among the curves, along with point and confidence interval estimates.

RESULTS

A total of 1189 patients from the UAB database had a $CD4^+$ cell count and CBC drawn on the same day. The baseline characteristics of these patients are listed in Table 1. The decision tree developed using the CART algorithm is

TABLE 1. Characteristics of the 1189 HIV-Infected Patients From UAB Included in the Study

Characteristic	Patients with CD4 Count ≤200 Cells/μL	Patients with CD4 Count >200 Cells/μL	Overall
Number (%)	381 (32%)	808 (68%)	1189
Demographics			
Median age (range)	38 (21–75)	39 (16–72)	39 (16–75)
Race			
White, N (%)	191 (50%)	463 (57%)	654 (55%)
African American, N (%)	179 (47%)	332 (41%)	511 (43%)
Other, N (%)	11 (3%)	13 (2%)	24 (2%)
Male, N (%)	314 (82%)	604 (75%)	918 (77%)
On HAART*, N (%)	271 (71%)	554 (69%)	825 (69%)
Median laboratory values (range)			
$CD4^+$ count, cells/μL†	69 (1–200)	467 (202–2243)	333 (1–2243)
Log viral load, copies/mL†	4.52 (1.26–6.45)‡	2.43 (1.41–6.00)§	2.97 (1.26–6.45)
TLC, cells/mL†	880 (40–3300)	1870 (0–7840)	1600 (0–7840)
Hemoglobin, g/dL†	12.8 (5.3–17.2)	14.2 (5.8–19.7)	13.8 (5.3–19.7)
Platelet count, $\times 10^3$ cells/mL†	185 (15–563)	219 (11–607)	210 (11–607)

*Highly active antiretroviral therapy, defined as ≥3 antiretroviral drugs within 30 days of the date of laboratory values.
†$P < 0.0001$ for difference between cohorts with CD4 count ≤200 cells/μL and CD4 count >200 cells/μL.
‡64 values missing.
§75 values missing.

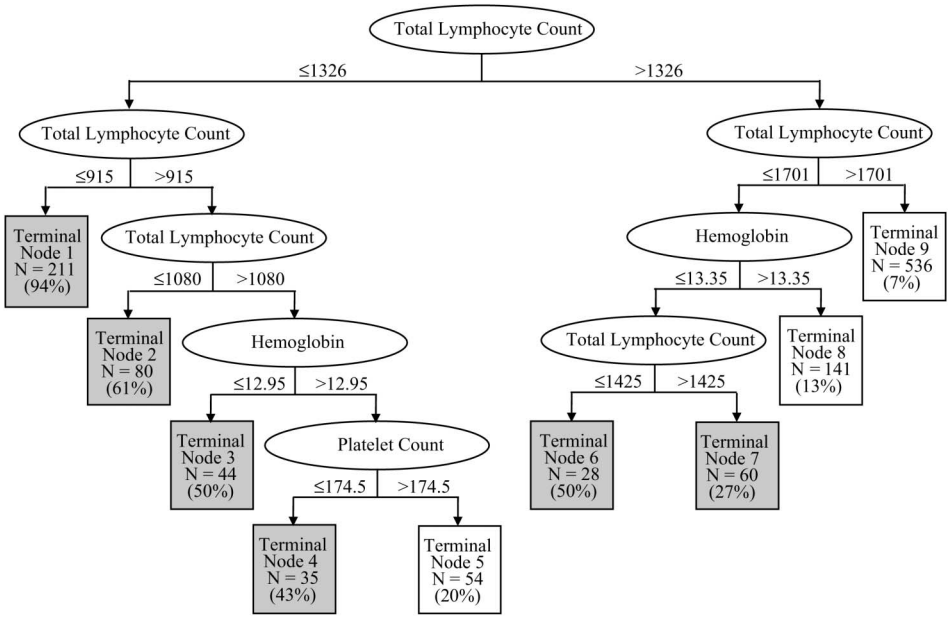

FIGURE 1. Decision tree generated by the classification and regression tree method based on UAB data. Using the results of the TLC, hemoglobin, and platelet count, one can follow the tree down step-by-step to a terminal node. The gray terminal nodes are classified as CD4 count ≤200 cells/μL, based on the ROC curve.

shown in Figure 1. Using data from a CBC, one can follow the decision tree down step-by-step to a terminal node and classification as a $CD4^+$ count <200 cells/μL or >200 cells/μL. A TLC of 1326 cells/mL is the first discriminating point with those above or below, going down separate sides of the tree. Each of the subsequent decision points is likewise based on one of the CBC laboratory values until a terminal node is reached. Each terminal node can be classified as greater than or less than a CD4 count of 200 cells/μL depending on how well that node identifies subjects and on unique local conditions, which determine whether test sensitivity or specificity is more important to emphasize. For this study, nodes are classified based on the best ROC curve, as defined by the curve with the largest AUC. Subjects ending up in a gray node (nodes 1–4, 6, and 7) are therefore classified as having a CD4 count ≤200 cells/μL. Although some of the terminal nodes with fewer subjects (eg, nodes 4 and 7) do not classify subjects extremely accurately, the UAB CART ROC curve as a whole (AUC = 0.888) correctly classified $CD4^+$ cell count significantly better than the TLC cut-point of 1200 cells/mL (AUC = 0.806; $P < 0.0001$; Table 2; Fig. 2A). Of note, the variables gender, BMI, and any antiretroviral therapy in the last 30 days (yes/no) were not discriminative of $CD4^+$ cell count.

The decision tree was then validated with data from 2 different populations. From the BVAMC, 512 sets of laboratory data (TLC, hemoglobin, platelet count, and $CD4^+$ cell count) were obtained from 204 HIV-infected patients, with a median $CD4^+$ count of 297 cells/μL (Table 3). Based on the UAB CART decision tree, the ROC curve generated had a significantly better AUC than the TLC cut-point of 1200 cells/mL (0.802 vs. 0.723, respectively; $P < 0.0001$). A CART decision tree developed directly from BVAMC data classified significantly better than the UAB CART decision tree, however, with an AUC of 0.886 ($P < 0.0001$; see Table 2; see Fig. 2B).

From Lusaka, Zambia, laboratory data were obtained from 596 HIV-infected women participating in a contraceptive clinical trial. The median $CD4^+$ count for these women was 471 cells/μL (see Table 3). Based on the UAB CART decision tree, the ROC curve generated had a significantly better AUC than the TLC cut-point of 1200 cells/mL (0.714 vs. 0.623, respectively; $P = 0.0009$). As with the BVAMC data, however, a CHAID decision tree developed directly from Zambian data classified significantly better than the UAB CART decision tree, with an AUC of 0.841 ($P < 0.0001$; see Table 2; see Fig. 2C).

DISCUSSION

In this study, we used a decision tree analysis to model whether the variables TLC, hemoglobin, platelet count,

TABLE 2. Characteristics of Different Laboratory Models to Identify $CD4^+$ Counts ≤200 Cells/μL

Model Classifying CD4 Count ≤200 Cells/μL	Sensitivity	Specificity	PPV	NPV	AUC
UAB					
TLC ≤1200 cells/mL	71%	90%	78%	87%	0.806
CART	83%	82%	69%	91%	0.888*
BVAMC					
TLC ≤1200 cells/mL	55%	89%	71%	81%	0.723
UAB CART	73%	76%	59%	86%	0.802*
BVAMC CART	81%	84%	71%	90%	0.886†
Zambia					
TLC ≤1200 cells/mL	27%	98%	59%	92%	0.623
UAB CART	52%	88%	32%	94%	0.714‡
Zambia CHAID	83%	68%	22%	97%	0.841†

*$P < 0.0001$ for the comparison with TLC ≤1200 cells/mL.
†$P \leq 0.0001$ for comparison with UAB CART.
‡$P = 0.0009$ for comparison with TLC ≤1200 cells/mL.

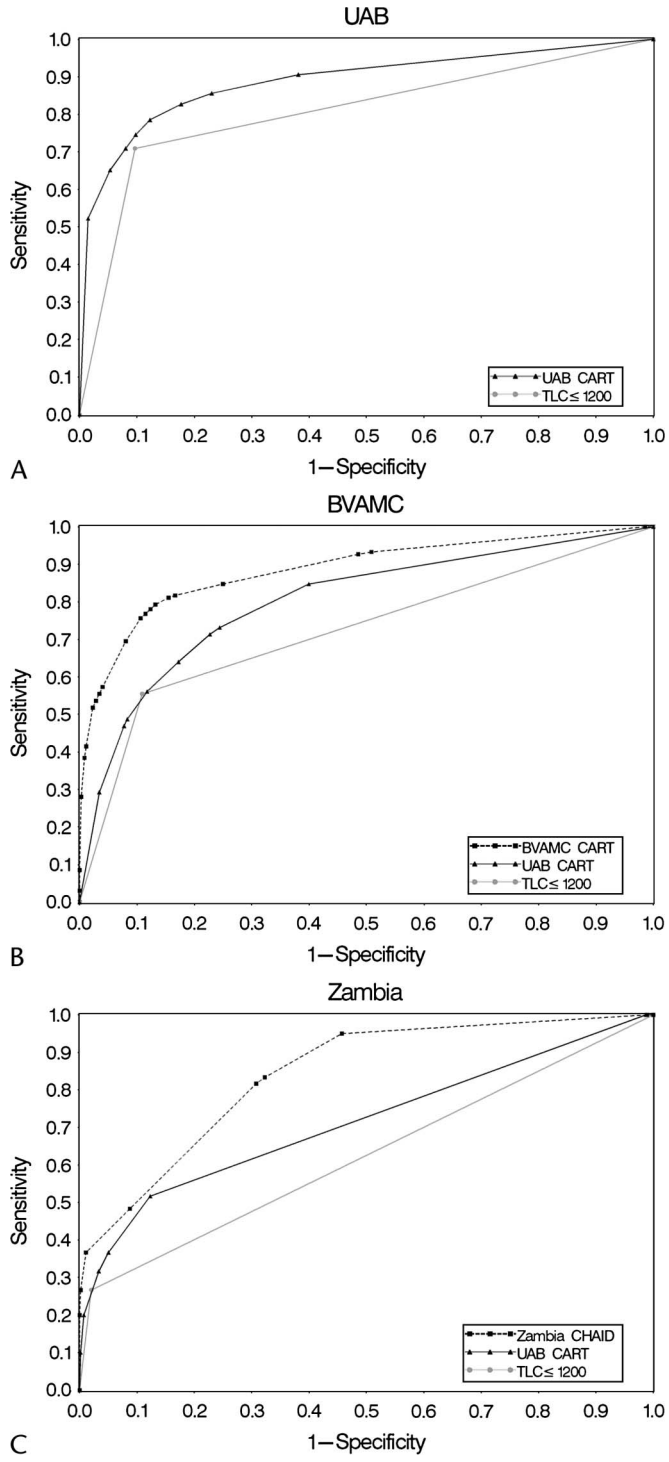

FIGURE 2. Comparison of ROC curves by different analysis techniques for the 3 different study cohorts. A, UAB data are used to compare the ROC curves from the TLC cut-point of 1200 cells/mL and the UAB CART decision tree. B, BVAMC data are used to compare the ROC curves from the TLC cut-point of 1200 cells/mL, the UAB CART decision tree, and the BVAMC CART decision tree. C, Zambian data are used to compare the ROC curves from the TLC cut-point of 1200 cells/mL, the UAB CART decision tree, and the Zambian CHAID decision tree.

TABLE 3. Median Laboratory Characteristics of the Different Cohorts

Laboratory Value (median)	UAB	BVAMC	Zambia
Number	1189	512	596
CD4 count, cells/μL (% <200 cells/μL)	333 (32%)	297 (32%)	471 (10%)
TLC, cells/mL	1600	1640	2270
Hemoglobin, g/dL	13.8	13.8	11.7
Platelet count, $\times 10^3$ cells/mL	210	221.5	247

gender, BMI, and any antiretroviral therapy within the previous 30 days (yes/no) identified CD4 counts ≤200/μL better than the TLC cut-point of 1200 cells/mL. Our model emphasizes the use of inexpensive, easily obtained variables that are relevant whether or not the patient is receiving antiretroviral medications. The variables TLC, hemoglobin, and platelet count were significant, and the UAB decision tree generated demonstrated an AUC significantly better than that of the TLC cut-point of 1200 cells/mL. When applied to data from other populations, although the UAB decision tree continued to perform significantly better than the TLC cut-point of 1200 cells/mL, a decision tree developed specifically from data from that population was clearly superior to both.

Decision trees have been used successfully to identify outcomes in medicine[36–38] but have not been used previously to classify $CD4^+$ cell counts. The primary benefit of using decision trees in rural settings is the ease with which they can be applied. Using the data from a patient's CBC to determine how to follow the tree down to the terminal node, any health care worker from a developing country can use the tree to classify a patient's CD4 status, and thus to make decisions about treatment with minimal training. In our study, using a decision tree algorithm incorporating TLC, hemoglobin, and platelet count to identify low $CD4^+$ cell counts was more effective than using the TLC cut-point of 1200 cells/mL in any single population.

When applied to a different population, however, both algorithms were inferior, based on the AUC, to a decision tree developed specifically from local data (see Table 2; see Fig. 2). The reasons why local data provide the best results may be related to the differences in baseline laboratory parameters. People in developing countries may have different "normal" baseline laboratory values because of local genetic, environmental, infectious, or nutritional factors that affect immunologic and hematologic parameters.[39–43] If, for this reason, locally developed decision trees determine low $CD4^+$ cell counts better for local populations, the same may apply to the WHO-recommended TLC cut-point of 1200 cells/mL. Our data support this with fairly different test characteristics between Birmingham and Lusaka for the TLC cut-point of 1200 cells/mL (see Table 2). This may also help to explain the differing results reported in the literature as to how well TLC classifies $CD4^+$ cell count.[4,5] Additional studies need to be done to confirm whether or not local models should be developed for different populations.

If a developing country site has additional resources, the local decision tree developed can be just a single step in a treatment algorithm. The Zambian decision tree, for example, has a 97% negative predictive value (NPV) but only a 22% positive predictive value (PPV; see Table 2). With such a high NPV, practitioners can be confident that patients who test "negative" (ie, have a $CD4^+$ count >200 cells/μL) actually have a $CD4^+$ count >200 cells/μL. With a low PPV, however, most patients who test "positive" (ie, have a $CD4^+$ count ≤ 200 cells/μL) actually have a $CD4^+$ count >200 cells/μL. Thus, patients who end up in a positive terminal node with a low PPV might be selected to receive further testing with a $CD4^+$ cell count if resources for only a limited number of $CD4^+$ cell count tests are available. For the Zambian data, 223 (37.4%) of 596 subjects tested positive, reducing the need for $CD4^+$ cell counts by more than 60%. This proportion would obviously be different for different populations, and treatment algorithms would need to be tailored to local situations.

Our study demonstrates that the discriminative ability of a decision tree model based on TLC, hemoglobin, and platelet count is significantly better than the WHO-recommended TLC cut-point of 1200 cells/mL. Furthermore, because the discriminative ability of both methods varies by population, a locally developed decision tree best identifies low $CD4^+$ cell counts. One limitation of our study is that we only examined 2 algorithms with 3 data sets. Whether or not a different algorithm based on 1 data set can be successfully applied to another population remains to be determined. At least 1 other study, however, found that a given model is not always applicable in another population.[44] If other studies confirm our result that the best model to identify a low $CD4^+$ cell count is one based on local data, an emphasis should be placed on encouraging local data analysis based on local treatment factors and priorities rather than on applying a single universal algorithm. In addition, continued progress must be made in identifying $CD4^+$ cell count assays that are affordable in resource-limited settings.[5,45]

REFERENCES

1. Kumarasamy N, Flanigan TP, Mahajan AP, et al. Monitoring HIV treatment in the developing world. *Lancet Infect Dis*. 2002;2:656–657.
2. Stephenson J. Cheaper HIV drugs for poor nations bring a new challenge: monitoring treatment. *JAMA*. 2002;288:151–153.
3. World Health Organization. *Scaling Up Antiretroviral Therapy in Resource-Limited Settings: Treatment Guidelines for a Public Health Approach, 2003 Revision*. Geneva: World Health Organization; 2004.
4. Schreibman T, Friedland G. Use of total lymphocyte count for monitoring response to antiretroviral therapy. *Clin Infect Dis*. 2004;38:257–262.
5. Crowe S, Turnbull S, Oelrichs R, et al. Monitoring of human immunodeficiency virus infection in resource-constrained countries. *Clin Infect Dis*. 2003;37(Suppl 1):S25–S35.
6. Montaner JS, Le TN, Le N, et al. Application of the World Health Organization system for HIV infection in a cohort of homosexual men in developing a prognostically meaningful staging system. *AIDS*. 1992;6: 719–724.
7. Blatt SP, Lucey CR, Butzin CA, et al. Total lymphocyte count as a predictor of absolute CD4+ count and CD4+ percentage in HIV-infected persons. *JAMA*. 1993;269:622–626.
8. Martin DJ, Sim JG, Sole GJ, et al. CD4+ lymphocyte count in African patients co-infected with HIV and tuberculosis. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1995;8:386–391.
9. Schechter MT, Le N, Craib KJ, et al. Use of the Markov model to estimate the waiting times in a modified WHO staging system for HIV infection. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1995;8: 474–479.
10. Beck EJ, Kupek EJ, Gompels MM, et al. Correlation between total and CD4 lymphocyte counts in HIV infection: not making the good an enemy of the not so perfect. *Int J STD AIDS*. 1996;7:422–428.
11. Wood R, Post F. Total lymphocyte count as a surrogate for CD4+ lymphocyte count in African patients coinfected with HIV and tuberculosis. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1996;11: 411–413.
12. Badri M, Wood R. Usefulness of total lymphocyte count in monitoring highly active antiretroviral therapy in resource-limited settings. *AIDS*. 2003;17:541–545.
13. Kumarasamy N, Mahajan AP, Flanigan TP, et al. Total lymphocyte count (TLC) is a useful tool for the timing of opportunistic infection prophylaxis in India and other resource-constrained countries. *J Acquir Immune Defic Syndr*. 2002;31:378–383.
14. Gange SJ, Lau B, Phair J, et al. Rapid declines in total lymphocyte count and hemoglobin in HIV infection begin at CD4 lymphocyte counts that justify antiretroviral therapy. *AIDS*. 2003;17:119–121.
15. Bedell R, Heath KV, Hogg RS, et al. Total lymphocyte count as a possible surrogate of CD4 cell count to prioritize eligibility for antiretroviral therapy among HIV-infected individuals in resource-limited settings. *Antivir Ther*. 2003;8:379–384.
16. Lee SS, Wong KH. The use of total lymphocyte count (TLC) as an independent criterion for initiating HAART in resource-poor countries. *J Infect*. 2005;50:66–67.
17. Stebbing J, Sawleshwarkar S, Michailidis C, et al. Assessment of the efficacy of total lymphocyte counts as predictors of AIDS defining infections in HIV-1 infected people. *Postgrad Med J*. 2005;81: 586–588.
18. van der Ryst E, Kotze M, Joubert G, et al. Correlation among total lymphocyte count, absolute CD4+ count, and CD4+ percentage in a group of HIV-1–infected South African patients. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1998;19:238–244.
19. Akanmu AS, Akinsete I, Eshofonie AO, et al. Absolute lymphocyte count as surrogate for CD4+ cell count in monitoring response to antiretroviral therapy. *Niger Postgrad Med J*. 2001;8:105–111.
20. Liotta G, Perno CF, Ceffa S, et al. Is total lymphocyte count a reliable predictor of the CD4 lymphocyte cell count in resource-limited settings? *AIDS*. 2004;18:1082–1083.
21. Akinola NO, Olasode O, Adediran IA, et al. 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: 579–581.
22. Kamya MR, Semitala FC, Quinn TC, et al. Total lymphocyte count of 1200 is not a sensitive predictor of CD4 lymphocyte count among patients with HIV disease in Kampala, Uganda. *Afr Health Sci*. 2004;4: 94–101.
23. Mocroft A, Kirk O, Barton SE, et al. Anaemia is an independent predictive marker for clinical prognosis in HIV-infected patients from across Europe. EuroSIDA Study Group. *AIDS*. 1999;13:943–950.
24. Glatt AE, Anand A. Thrombocytopenia in patients infected with human immunodeficiency virus: treatment update. *Clin Infect Dis*. 1995;21: 415–423.
25. Wheeler DA, Gibert CL, Launer CA, et al. Weight loss as a predictor of survival and disease progression in HIV infection. Terry Beirn Community Programs for Clinical Research on AIDS. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1998;18:80–85.
26. Maas JJ, Dukers N, Krol A, et al. Body mass index course in asymptomatic HIV-infected homosexual men and the predictive value of a decrease of body mass index for progression to AIDS. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1998;19:254–259.
27. Lau B, Gange SJ, Phair JP, et al. Rapid declines in total lymphocyte counts and hemoglobin concentration prior to AIDS among HIV-1-infected men. *AIDS*. 2003;17:2035–2044.
28. Spacek LA, Griswold M, Quinn TC, et al. Total lymphocyte count and hemoglobin combined in an algorithm to initiate the use of highly active antiretroviral therapy in resource-limited settings. *AIDS*. 2003;17: 1311–1317.
29. Schechter M, Zajdenverg R, Machado LL, et al. Predicting CD4 counts in HIV-infected Brazilian individuals: a model based on the World Health

Organization staging system. *J Acquir Immune Defic Syndr*. 1994;7: 163–168.
30. Mwamburi DM, Ghosh M, Fauntleroy J, et al. Predicting CD4 count using total lymphocyte count: a sustainable tool for clinical decisions during HAART use. *Am J Trop Med Hyg*. 2005;73:58–62.
31. Costello C, Nelson KE, Jamieson DJ, et al. Predictors of low CD4 count in resource-limited settings: based on an antiretroviral-naive heterosexual Thai population. *J Acquir Immune Defic Syndr*. 2005;39:242–248.
32. Zhang HP, Singer B. *Recursive Partitioning in the Health Sciences*. New York: Springer; 1999.
33. Kass GV. An exploratory technique for investigating large quantities of categorical data. *Appl Stat*. 1980;29:119–127.
34. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–845.
35. De Long DM. Sample 520: nonparametric comparison of areas under correlated ROC curves. Available at: http://support.sas.com/ctx/samples/index.jsp?sid=520. Accessed April 6, 2006.
36. Kammerer JS, McNabb SJ, Becerra JE, et al. Tuberculosis transmission in nontraditional settings: a decision-tree approach. *Am J Prev Med*. 2005; 28:201–207.
37. Cancre N, Bois F, Gresenguet G, et al. Screening blood donations for hepatitis C in Central Africa: analysis of a risk- and cost-based decision tree. *Med Decis Making*. 1999;19:296–306.
38. El-Solh A, Mylotte J, Sherif S, et al. Validity of a decision tree for predicting active pulmonary tuberculosis. *Am J Respir Crit Care Med*. 1997;155:1711–1716.
39. Fakhir S, Ahmad P, Faridi MA, et al. Cell-mediated immune responses in malnourished host. *J Trop Pediatr*. 1989;35:175–178.
40. Ozkan H, Olgun N, Sasmaz E, et al. Nutrition, immunity and infections: T lymphocyte subpopulations in protein-energy malnutrition. *J Trop Pediatr*. 1993;39:257–260.
41. Greenberg PL, Gordeuk V, Issaragrisil S, et al. Major hematologic diseases in the developing world—new aspects of diagnosis and management of thalassemia, malarial anemia, and acute leukemia. *Hematology (Am Soc Hematol Educ Program)*. 2001;479–498.
42. Yip R. Iron deficiency: contemporary scientific issues and international programmatic approaches. *J Nutr*. 1994;124(8 Suppl):1479S–1490S.
43. Diallo D, Tchernia G. Sickle cell disease in Africa. *Curr Opin Hematol*. 2002;9:111–116.
44. Ferris DC, Dawood H, Magula NP, et al. Application of an algorithm to predict CD4 lymphocyte count below 200 cells/mm(3) in HIV-infected patients in South Africa. *AIDS*. 2004;18:1481–1482.
45. Vermund SH, Powderly WG. Developing a human immunodeficiency virus/acquired immunodeficiency syndrome therapeutic research agenda for resource-limited countries: a consensus statement. *Clin Infect Dis*. 2003;37(Suppl 1):S4–S12.

