The absolute CD4 cell count is used routinely in the evaluation and monitoring of HIV-infected people; in determining eligibility for trials of new anti-retroviral therapies and as an endpoint measure of drug efficacy. In addition, since January 1993, a CD4 count of less than 200 cells/mm³ has been incorporated into the Centers for Diseases Control (CDC) AIDS case definition. However, imprecision of CD4 cell testing is considerable, even among laboratories participating in a quality control programme. The absolute CD4 cell count is calculated from both the total and differential white blood cell count. Variability in the CD4 measurement may occur as the result of laboratory test error, as well as intraperson temporal fluctuations in these measures due to biological factors such as diurnal variation, stress and acute infections. Since certain management decisions, such as initiation of prophylaxis against Pneumocystis carinii pneumonia (PCP), may be based on a single CD4 cell count, CD4 measurement error may have important clinical consequences. Hoover has demonstrated how patients with a true CD4 cell count far above a given CD4 threshold may test below it with a high probability as a result of CD4 measurement error, unless confirmatory retests are used, and both these tests are below the threshold limit.

In contrast, the CD4% which is measured directly on a flow cytometer is much less variable, with a coefficient of variability of 6–24% compared with 19–40% for the absolute CD4 cell count. Several studies have also shown that the CD4% (or the rate of change in CD4%) is a better predictor of clinical progression. It has therefore been suggested that the CD4% rather than the absolute cell number should be used for patient monitoring. A further advantage of using CD4% is financial, since it requires only the flow cytometer results, and not the white blood cell count and differential.

An important prerequisite to the more widespread adoption of CD4% measurements into clinical practice and research, is a demonstration of the statistical relationship between the absolute CD4 count and CD4%.
and the influence on this relationship of important demographic and clinical factors, such as gender, risk group, age, an AIDS diagnosis, and use of zidovudine therapy and/or PCP prophylaxis.

METHODS

Study Population
Patients eligible for analysis were 1017 HIV-1 seropositive subjects attending one of two hospital-based HIV clinics in West London who were: (i) initially AIDS free and zidovudine-naive, and (ii) had at least one paired CD4 count and CD4% measurement during follow-up. The patient's clinical status during follow-up was categorized according to the Centers for Disease Control 1987 classification. Linked pharmacy records provided information on date of initiation of zidovudine and PCP prophylaxis (oral trimethoprim-sulphamethoxazole, dapsone, or aerosolized pentamidine).

Measurement of CD4 Count and CD4%
All patients were reviewed periodically and their T-lymphocyte subsets measured as clinically indicated by flow cytometry (FACScan, Becton-Dickinson, Cowley, Oxford). The median interval between CD4 count measurements was 5 weeks but ranged from weekly to 5-monthly, and the median number of paired observations per patient was six. Absolute CD4 lymphocyte count was calculated from the product of the white cell count, the lymphocyte percentage, and the CD4%. All CD4 measurements were made using the same flow cytometer, sample preparation and haematology laboratory over the duration of the study.

Statistical Approach
The CD4 cell count and CD4% measurements were log transformed to make the data more normally distributed for further analysis. The relationship between CD4 count and CD4% was then explored using a two-level modelling approach (ML3: Software for 3-level analysis) which is an extension of ordinary multiple regression. In the typical multiple regression model, individual observations of the dependent variable are assumed to be sampled independently of each other, so that the random errors are uncorrelated. In our data, this assumption does not hold because patient-specific effects cause CD4% from the same patient to be correlated. The two-level model formally identifies these patient-specific effects as additional sources of random variation at the patient to patient level. In a preliminary analysis, a likelihood ratio test demonstrated significant between-patient heterogeneity ($P < 0.001$), suggesting the need for a model that allowed for patient differences.

RESULTS

Study Population
Table 1 shows the clinical characteristics of the study population and the overall distribution of their serial CD4 cell measurements. The majority of the 1017 patients were male (94%) and their average age at analysis was 34.8 years (range 16–66 years). The risk group status of the participants was as follows: 956 (94%) homosexual/bisexual; 35 (3%) injecting drug users; and 26 (3%) heterosexual. During follow-up, 572 (56%) patients received both zidovudine (ZDV) treatment and PCP prophylaxis (trimethoprim-sulphamethoxazole, aerosolized pentamidine or dapsone); 219 (22%) received ZDV alone, 69 (7%) PCP prophylaxis alone, and 157 (15%) neither therapy. Of the 1017 patients, in 600 (60%), all their available serial CD4 count measurements were below 200 cells/mm$^3$, while in 90 (10%) all their measurements were above 500 cells/mm$^3$.

Estimated CD4 Cell Count from CD4%
9203 paired serial measurements of CD4 count and CD4% from 1017 HIV-1 seropositive patients were available
for analysis. The median number of paired observations per patient was 6 (range 1–45). The standard deviation of the log CD4 cell count and log CD4% in this patient sample was 1.38 and 0.94, respectively, and as expected, the correlation between log CD4 cell count and log CD4% was high at 0.88.

A strong linear relationship was found between the dependent variable, log CD4 and the independent variable, log CD4% as represented by:

\[
\log (\text{CD4 count}) = 1.78 + 1.26 \log (\text{CD4 %})
\]

where 1.78 and 1.26 represent the estimated intercept and slope of the model with standard deviations of 0.49 and 0.22, respectively.

Figure 1 is a diagrammatic representation of the predicted CD4 cell count and 95% CI from a given CD4%, based on this relationship. The precision of the CD4 cell count estimates is markedly reduced at higher CD4% values, as shown by the diverging 95% confidence boundaries. At a given CD4% of 5%, 15% and 30%, the corresponding estimated CD4 cell count and 95% CI are 45 cells/mm³ (17–117 cells/mm³), 182 cells/mm³ (64–499 cells/mm³), and 438 cells/mm³ (132–1395 cells/mm³), respectively.

**Influence of Covariates**

Table 2 shows the relative influence of selected demographic and clinical variables on the relationship between log CD4 count and log CD4%. The percentage change (increase or decrease) in the CD4 count estimated from the CD4%, after adjusting for a particular variable, was calculated from the exponent of the covariate coefficient. For example, when compared to homosexual men, the predicted CD4 cell count for a given CD4% was lower among heterosexuals by 30%, \(1 - e^{-0.35}\), and among injecting drug users by 17%, \(1 - e^{-0.19}\). Similarly, the estimated CD4 count was lower following an AIDS diagnosis by 21%, \(1 - e^{-0.24}\) and after initiation of ZDV therapy and PCP prophylaxis by 19%, \(1 - e^{-0.21}\) and 10%, \(1 - e^{-0.10}\), respectively. Neither gender nor age had any significant impact on this relationship.

The estimated CD4 cell count and 95% CI for three levels of CD4% (5%, 15% and 30%), are presented in Table 3, stratified by risk group status (homosexual/bisexual, injecting drug user and heterosexual); presence or absence of an AIDS diagnosis, and use of ZDV and/or PCP prophylaxis. The Table shows more clearly the separate and combined effects of risk group, clinical status and therapy on the estimated CD4 cell count at a low, medium level, and high CD4%.

Figure 1. Graphical representation of estimated CD4 cell count and 95% confidence intervals from CD4%

### Table 2: Multiple regression analysis using 2-level modelling of log CD4 count on log CD4%

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercepts</td>
<td>2.23</td>
<td>0.025</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>log CD4%</td>
<td>1.12</td>
<td>0.020</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Homosexual/bisexual (ref)*</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Heterosexual</td>
<td>–0.35</td>
<td>0.110</td>
<td>0.0014</td>
</tr>
<tr>
<td>Injecting drug user</td>
<td>–0.19</td>
<td>0.094</td>
<td>0.043</td>
</tr>
<tr>
<td>AIDS diagnosis</td>
<td>–0.24</td>
<td>0.016</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>ZDV therapy</td>
<td>–0.21</td>
<td>0.019</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>PCP prophylaxis</td>
<td>–0.10</td>
<td>0.016</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>ZDV × log CD4% (interaction)</td>
<td>0.07</td>
<td>0.018</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* Age and gender had no significant influence on the model (\(P = 0.285\) and 0.484, respectively).

b Variance of the intercept and log CD4% is 0.225 and 0.053, respectively, and their corresponding covariance is 0.012.
greater extent among heterosexual risk groups relative to homosexual men.

We sought to explain the slightly lower predicted CD4 counts among users of ZDV and/or PCP prophylaxis by the effect of these therapies (mainly trimethoprim-sulphamethoxazole and zidovudine) on the total white cell count. Based on a multi-level regression analysis, the serial white cell counts of patients after commencing PCP prophylaxis were significantly lower than the white counts of patients who had never received prophylaxis ($P = 0.0001$). Similarly, there was a trend towards a lower white cell count following initiation of ZDV compared to a similar group of patients who had never received ZDV therapy ($P = 0.08$), matched on time from infection. However, addition of white cell count to the full statistical model did not substantially alter the impact of ZDV therapy or PCP prophylaxis on the relationship between CD4 and CD4%.

**CONCLUSIONS**

Previous studies have reported that a CD4% of 10%, 20%, 25% and 30% corresponds broadly to a CD4 cell count of 100, 200, 250 and 300 cells/mm$^3$ respectively, for the purpose of clinical prediction.$^{13,14}$ Indeed, the Centers for Disease Control (CDC) used a different formula in their revised classification system for HIV infection, where a cutoff of 14% and 29% was considered equivalent to a CD4 count of 200 cells/mm$^3$ and 500 cells/mm$^3$, respectively.$^{15}$ While at the lower range of CD4% (5–10%), our findings are consistent with these previous studies, (i.e. 5% = 45 CD4 cells/mm$^3$, 10% = 109 CD4 cells/mm$^3$), at higher levels of CD4%, it is clear that this represents an oversimplification of the relationship between CD4 count and CD4%.

For example, in our study population, a CD4% of 20% was found to correspond to a CD4 cell count of 262 rather than 200 cells/mm$^3$, and a CD4% of 30% equivalent to a CD4 cell count of 438 rather than 300 cells/mm$^3$. Our model shows reasonable agreement with three other models derived from an analysis of 1000 individuals each at the University of Washington (UW) ($\log CD4 = 1.96 + 1.29 (\log CD4%)$), Maryland Medical Laboratory (MML) ($\log CD4 = 2.14 + 1.24 (\log CD4%)$), and Rush-Presbyterian–St Luke’s Medical Center (RUSH) ($\log CD4 = 1.81 + 1.28 (\log CD4%)$) in the US.$^{16}$ Table 4 presents a comparison of the predicted CD4 cell values at given levels of CD4% for these three models compared to our own findings from two centres in London (LOND). It is noteworthy that the predicted CD4 cell count from our model agrees particularly well with the model from the Rush-Presbyterian Center, but is consistently lower than the values predicted from the other two US models, although their estimated CD4 values

<table>
<thead>
<tr>
<th>Risk group</th>
<th>Homosexual/bisexual</th>
<th>Injecting drug user</th>
<th>Heterosexual</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4%</td>
<td>5% (CI)</td>
<td>15% (CI)</td>
<td>30% (CI)</td>
</tr>
<tr>
<td>AIDS-free:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZDV only</td>
<td>51 (17–158)</td>
<td>438 (79–2424)</td>
<td>362 (65–2006)</td>
</tr>
<tr>
<td>PCP prophylaxis only</td>
<td>51 (17–156)</td>
<td>381 (69–2110)</td>
<td>316 (57–1747)</td>
</tr>
<tr>
<td>Both ZDV and</td>
<td>47 (15–143)</td>
<td>396 (72–2193)</td>
<td>328 (59–1816)</td>
</tr>
<tr>
<td>PCP prophylaxis</td>
<td>56 (15–183)</td>
<td>421 (76–2332)</td>
<td>349 (63–1930)</td>
</tr>
<tr>
<td>Neither therapy</td>
<td>56 (15–173)</td>
<td>421 (76–2332)</td>
<td>349 (63–1930)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Risk group</th>
<th>Homosexual/bisexual</th>
<th>Injecting drug user</th>
<th>Heterosexual</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIDS diagnosis:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZDV only</td>
<td>41 (13–124)</td>
<td>346 (62–1913)</td>
<td>286 (52–1583)</td>
</tr>
<tr>
<td>PCP prophylaxis only</td>
<td>41 (13–123)</td>
<td>301 (54–1665)</td>
<td>249 (45–1378)</td>
</tr>
<tr>
<td>Both ZDV and</td>
<td>37 (12–112)</td>
<td>313 (56–1731)</td>
<td>259 (47–1433)</td>
</tr>
<tr>
<td>PCP prophylaxis</td>
<td>37 (12–112)</td>
<td>313 (56–1731)</td>
<td>259 (47–1433)</td>
</tr>
<tr>
<td>Neither therapy</td>
<td>45 (15–136)</td>
<td>332 (60–1840)</td>
<td>275 (50–1523)</td>
</tr>
</tbody>
</table>

**Table 3 Estimated CD4 count (cells/mm$^3$) and 95% confidence interval (CI) for CD4% at 5%, 15% and 30%, according to risk group, clinical status, and use of zidovudine (ZDV) and/or PCP prophylaxis**
all lie within our wide CI. As might be expected, from the diverging CI at high CD4 counts, the disparity between the models is greatest with higher CD4% values. Thus a CD4% of 5% corresponds with a predicted absolute CD4 cell count of 45 cells/mm³ in our London cohort, as compared with 57 (UW), 62 (MML) and 46 (RUSH) cells/mm³. Similarly, a CD4% of 30%, corresponds to 438 (LOND), 571 (UW), 577 (MML) and 475 (RUSH) cells/mm³. The most likely explanation for this discrepancy is the much greater representation of patients with higher CD4% and CD4 cell counts in the US study population groups, resulting in more stable and precise estimates at high CD4% than in our own cohort. In addition, the analytic approaches used in modelling the relationship differ between the studies. While the three US models used cross-sectional data based on a single paired CD4 measurement of CD4 count and CD4% in 1000 individuals, in our own cohort, a median of six paired serial measurements from 1017 individuals were analysed, allowing us to incorporate within-person heterogeneity into the model. Finally, measurement error due to technical differences in sample collection and processing at the various clinical sites may have also contributed to the observed disparity between the models.

The nomogram (Figure 1) should be useful to the clinician and researcher in translating values of a given or observed CD4% to approximate absolute CD4 cell numbers, particularly when the percentage may be the only laboratory information available. A similar previous study demonstrated the usefulness of the total lymphocyte count as a substitute for the CD4 cell count or percentage, when these measures are unavailable.17 It is noteworthy that the intervals presented in Figure 1 are CI that relate to the patient’s underlying CD4 count, rather than prediction intervals for a single CD4 measurement, which is less relevant to the clinical picture. However, the wide CI for the estimated CD4 cell counts observed in our analysis, particularly with high CD4 cell values, limits the practical application of this predictive model.

The relationship between CD4 count and CD4% is further complicated by the marked effect of certain covariates. For patients from non-homosexual risk groups, following an AIDS diagnosis, or after initiation of ZDV or PCP prophylaxis, further downward adjustments of the estimated CD4 count for a given CD4% are necessary. These findings are only in part explained by the tendency towards a lower white cell count following initiation of ZDV therapy, or with an AIDS diagnosis.

Before the more widespread adoption into clinical practice of CD4% in preference to CD4 count, further studies are needed both to: refine the precision of the estimated CD4 cell count at high CD4% value; to quantify the measurement error associated with CD4%; to determine the basis for the lower CD4 counts in non-homosexual risk groups, following an AIDS diagnosis and with anti-retroviral therapy; and to examine the relative prognostic value of CD4% versus CD4 count in other population subgroups.

REFERENCES
8 Centers for Disease Control: Revision of the CDC surveillance case definition for acquired immunodeficiency syndrome. MMWR 1987. 36 (Suppl 25): 1S–15S.
with zero means and variances on both random variables and assumed to be normally distributed.

\[ Y_{ij} = \alpha_j + \beta X_{ij} + \varepsilon_{ij} \]  

where \( \alpha \) is the regression intercept for the \( j \)th patient and \( \beta \) is the regression slope of log CD4% which is assumed to be constant across all patients (i.e., no allowance for between patient variance). \( \varepsilon_{ij} \) is the residual for the \( i \)th CD4 count in the \( j \)th patient and is assumed to be normally distributed with zero mean and constant variance \( \sigma_w^2 \). To make allowance for the between-patient variation (two-level model approach), the relationship between log CD4 count and log CD4% was further modified:

\[ Y_{ij} = (\alpha + u_j) + (\beta + v_j) X_{ij} + \varepsilon_{ij} \]  

where \( u_j \) is the regression intercept for the \( j \)th patient and is assumed to be normally distributed with zero mean and constant variance for between patient variance.

Equation (2) can be re-expressed as:

\[ Y_{ij} = (\alpha + u_j) + (\beta + v_j) X_{ij} + \varepsilon_{ij} \]  

where \( u_j \) is the regression intercept for the \( j \)th patient and \( v_j \) is the residual for the \( i \)th CD4 count in the \( j \)th patient. Equation (2) can be re-expressed as:

\[ Y_{ij} = (\alpha + u_j) + (\beta + v_j) X_{ij} + \varepsilon_{ij} \]  

To which \( j \)th patient departs from the average, and are both random variables and assumed to be normally distributed with zero means and variances \( \sigma_u^2 \) and \( \sigma_v^2 \), respectively. The covariance between \( u_j \) and \( v_j \) is denoted by \( \sigma_{uv} \).

The 95% confidence interval of log CD4 count for a given value of log CD4% is estimated from:

\[ \text{predicted } \bar{Y} \pm 1.96 \sqrt{\sigma^2_b} \]  

where \( \sigma_b^2 \) is the between-patient variance and is defined as:

\[ \sigma^2_b = \sigma^2_\alpha + 2 \sigma_{\alpha \beta} X_{ij} + \sigma^2_\beta X^2_{ij} \]  

This formula does not explicitly allow for the variance in the estimation of \( \alpha \) and \( \beta \), but this is small in comparison with the terms \( \sigma^2_\alpha \), \( \sigma^2_\beta \) and \( \sigma_{\alpha \beta} \).

Testing for Between-Patient Heterogeneity

A likelihood ratio test was used to test for the presence of between-patient heterogeneity in the relationship between CD4 count and CD4%.

If \( L_1 \) is the likelihood of the above model (2) (i.e., allowing for between-patient variation) and \( L_2 \) is the likelihood of the model that does not allow for any between-patient variation, then the likelihood ratio test is represented by:

\[ -2[\ln(L_1) - \ln(L_2)] = \chi^2_{k-p} \]

where \( k \) and \( p \) represent the respective degrees of freedom for the two models.

Effect of Covariates

To examine for the influence of selected covariates, (2) is extended as shown below:

\[ Y_{ij} = \alpha_j + \beta_j X_{ij} + \sum_{h=1}^{n} C_h Z_{ijh} + \varepsilon_{ij} \]  

where \( n \) is the number of covariates of interest, \( Z_{ijh} \) is the \( h \)th covariate of interest for the \( j \)th patient, and \( C_h \) is the corresponding coefficient.