CLINICAL AND DIAGNOSTIC LABORATORY IMMUNOLOGY, July 1995, p. 462–468 Vol. 2, No. 4
1071-412X/95/$04.00+0
Copyright © 1995, American Society for Microbiology

Interlaboratory Variability of CD8 Subset Measurements by Flow Cytometry and Its Applications to Multicenter Clinical Trials

A. L. LANDAY,[1]* D. BRAMBILLA,[2] J. PITT,[3] G. HILLYER,[4] D. GOLENBOCK,[5] J. MOYE,[6] S. LANDESMAN,[7] J. KAGAN,[6] AND THE NIAID/NICHD WOMEN AND INFANTS TRANSMISSION STUDY GROUP†

Rush-Presbyterian-St. Luke's Medical Center, Chicago, Illinois[1]; New England Research Institute, Watertown,[2] and Boston City Hospital, Boston,[5] Massachusetts; Columbia Presbyterian Hospital, New York,[3] and State University of New York, Brooklyn,[7] New York; University of Puerto Rico, San Juan, Puerto Rico[4]; and National Institutes of Health, Rockville, Maryland[6]

Received 9 December 1994/Returned for modification 27 January 1995/Accepted 1 May 1995

Recent studies have demonstrated the utility of measuring subsets of CD8$^+$ T cells as prognostic markers in epidemiology cohort studies of human immunodeficiency virus (HIV)-infected patients. Most of these studies evaluating the value of CD8$^+$ T-cell subsets have been performed at single centers, and few data are available on variability in the measurement of the CD8$^+$ cell populations in multicenter trials. In the current study, we addressed this question by evaluating interlaboratory variability from the five laboratories enrolled in the Women and Infants Transmission Study sponsored by the National Institutes of Health. This study evaluated 35 HIV-positive and 28 HIV-negative proficiency testing samples sent to the laboratories for evaluation. The study focused on the robust coefficient of variation (RCV) for CD38 (11%), HLA-DR (21%), and CD57 (15%) expression on the CD8$^+$ population. Data from the current study indicated that the variability in these measurements is greater than that for CD3$^+$ CD4$^+$ (RCV, 5%) and CD3$^+$ CD8$^+$ (RCV, 5%) cells. Knowledge of the variability of the CD8$^+$ subset measurements should guide investigators in the design and analysis of clinical trials and epidemiology studies. Ability to obtain improved interlaboratory agreement on CD8$^+$ subset measurements will facilitate further evaluation of these markers in HIV studies.

CD4$^+$ T-cell depletion is considered to be the hallmark of human immunodeficiency virus (HIV) disease (20). More recent studies have shown profound changes in absolute numbers (and proportions) and function of CD8$^+$ T cells, and their subsets are prevalent in HIV-infected individuals across the spectrum of HIV disease (1, 15, 19, 22, 23, 27, 32, 37). For example, CD8$^+$ T cells have been shown to mediate HIV-specific cytolytic T-cell killing in vitro (7, 17, 18, 31, 38). With disease progression, both the frequency and function of HIV-specific CD8$^+$ cytolytic T cells decline (25, 27) and this is seen to parallel an increase in the HIV burden as measured both by proviral DNA in peripheral blood CD4$^+$ T cells and the levels of infectious virus in plasma and peripheral blood mononuclear cells (30). Levy et al. (24, 39–41) have shown that, in vitro, CD8$^+$ T cells can inhibit HIV replication in CD4$^+$ cells. Perhaps the best evidence that CD8$^+$ T cells have a role in controlling HIV replication is the fact that in acute HIV infection, the appearance of CD8$^+$ cytolytic T cells immediately precedes the precipitous drop in the HIV load in the periphery whereas the appearance of specific anti-HIV neutralizing antibodies follows the decline in the viral burden (2, 5, 9, 30).

Although CD8$^+$ T-cell measurements are a standard component in the most commonly utilized immunophenotyping guidelines for HIV-positive individuals (6, 8, 26), attention to CD8$^+$ T-cell measurements in the settings of patient care, clinical trials, and epidemiologic studies has not been strong because of several factors. First, CD8$^+$ T-cell numbers (and percentages) have not demonstrated the strong prognostic power for disease progression and/or survival that has been seen with CD4. Second, changes in CD8$^+$ T cells have not been seen to reflect antiviral drug activity or correlate with therapeutic efficacy and therefore have offered little information of practical value to clinical investigators and primary care providers. In fact, the reasons most often cited for inclusion of CD8$^+$ T cells in an immunophenotyping panel is to obtain a CD4/CD8 ratio and for quality control (8). However, over the past several years, studies have shown that specific subsets of CD8$^+$ T cells in HIV-positive individuals, most notably, those positive for expression of the CD38 and DR markers, have significant predictive value for disease progression (16, 32). Unfortunately, this has not been widely appreciated, probably because of the unfounded presumption that the lack of prognostic power for the "total" CD8$^+$ T-cell value would preclude any prognostic value for subsets of CD8$^+$ T cells. Another factor which has likely contributed to the lack of incorporation

* Corresponding author. Mailing address: Department of Immunology/Microbiology, Rush-Presbyterian-St. Luke's Medical Center, 1653 West Congress Parkway, Chicago, IL 60612. Phone: (312) 942-2849. Fax: (312) 942-6787.

† Principal investigators were Ruth Tuomala, Brigham and Women's Hospital, Boston, Mass.; Harold Fox and Jane Pitt, Columbia Presbyterian Hospital, New York, N.Y.; Sheldon Landesman and Hermann Mendez, State University of New York, Brooklyn; Clemente Diaz, University of Puerto Rico, San Juan; Kenneth Rich, University of Illinois at Chicago, Chicago, Ill.; and Sonja McKinlay, New England Research Institute, Watertown, Mass. Study coordinators were Donna Mesthene, Brigham and Women's Hospital; Alice Higgins, Columbia Presbyterian Hospital; Gail Moroso, State University of New York, Brooklyn; Edna Pacheco-Acosta, University of Puerto Rico; and Geraldine Alexander, University of Illinois at Chicago. Program officers were Mary Glenn Fowler and Amy R. Sheon, National Institute of Allergy and Infectious Diseases, Bethesda, Md., and Pamela Stratton and Anne Willoughby, National Institute of Child Health and Human Development, Bethesda, Md.

of these CD8$^+$ T-cell subsets in clinical practice is the lack of data addressing the variability in the measurement of these lymphocyte subpopulations. The work described herein represents the first investigation of interlaboratory variability in the enumeration of the CD38- and HLA-DR-expressing subsets of the CD8$^+$ T-cell population. The results of these studies indicate that the interlaboratory measurement variability of these markers is greater than that of both CD3$^+$ CD4$^+$ and CD3$^+$ CD8$^+$ T-cell measurements. This information will help in sample size calculations for AIDS clinical trials and epidemiology studies in which these measures serve as important variables. Furthermore, these data suggest a need for better standardization and quality control of these markers.

MATERIALS AND METHODS

Division of AIDS flow cytometry proficiency testing program. The flow cytometry proficiency testing program of the Division of AIDS of the National Institutes of Health consists of the monthly evaluation of three whole blood samples obtained from either HIV-positive or HIV-negative donors. Whole blood samples are obtained in heparin anticoagulant and shipped by a commercial carrier for overnight delivery to the participating laboratories. In most cases, the samples are received by the laboratories within 18 h after being drawn. Each of the participating laboratories is required to analyze the samples by utilizing a standard monoclonal antibody panel, as well as protocol-specific panels for the studies in which they participate. Analysis is conducted in accordance with guidelines established by the National Institute of Allergy and Infectious Diseases Division of AIDS Flow Cytometry Advisory Committee (6). These guidelines require that samples be analyzed by utilizing forward angle and 90° scatter gating with gate purity and recovery determined by the CD45/CD14 reagent. Data reported in the proficiency testing program are adjusted according to the percentage of lymphocytes within the analysis region as determined by the CD45/CD14 reagent. The laboratories are subject to certification on the basis of differences between their reported CD4 percentages and the median values obtained for the entire group. Certification status is ranked as certified, provisionally certified, probationary, or suspended. Each of the sites participating in this study maintained a certified status throughout the study period of September 1991 to December 1993.

WITS. The proficiency testing data evaluated for this study were obtained from the five sites in Women and Infants Transmission Study I (WITS I). WITS I (1990 to 1994) was a multicenter epidemiologic study that evaluated the natural history of HIV infection in pregnant mothers and their infants, including factors related to perinatal transmission. The study sites were located at the University of Illinois at Chicago; the University of Puerto Rico, San Juan; Columbia Presbyterian Hospital, New York, N.Y.; the State University of New York Health Science Center, Brooklyn; and a consortium of sites in Boston (Maxwell Finland Laboratory, Boston City Hospital) and Worcester, Mass. Each of the WITS sites was equipped with a FACScan flow cytometer (Becton-Dickinson Corporation, San Jose, Calif.) and participated in the proficiency testing program described above.

Monoclonal antibody panel. A standardized panel of monoclonal antibody reagents was utilized in evaluating HIV-positive pregnant women and their infants enrolled in the WITS cohort. The monoclonal antibody panel consisted of the following reagents: isotype control (immunoglobulins G1 and G2a), CD45/CD14 (gating reagent), CD4/CD8 (compensation reagent), CD3/CD4 (CD4$^+$ T cells), CD3/CD8 (CD8$^+$ T cells), CD3/CD19 (T and B cells), CD3/CD16, CD56 (T and NK cells), CD8/CD38 (activated CD8 subset), CD8/HLA-DR (activated CD8 subset), and CD57/CD8 (functional CD8 subset). Each of the monoclonal antibody pairs was obtained as fluorescein isothiocyanate- and phycoerythrin-conjugated antibodies, respectively, and prepared as standardized reagents through Becton-Dickinson Corporation. Each site received the same lot of reagents for evaluation of proficiency testing and patient samples.

Sample preparation and flow cytometric analysis. A whole-blood lysing procedure was utilized for preparation of the samples as described by the Division of AIDS flow cytometry guidelines (6). Briefly, 100 μl of the sample was aliquoted into a 5-ml polystyrene tube, 20 μl of the appropriate monoclonal antibody was added, and the mixture was incubated for 15 min at room temperature. Following the incubation period, 2 ml of FACSLyse reagent was added to lyse the erythrocytes and the samples were incubated for an additional 10 min at room temperature. The samples were centrifuged at 1,500 rpm for 5 min to pellet the cells. Each sample was washed with 1 ml of phosphate-buffered saline containing 3% fetal calf serum and 0.1% sodium azide and then fixed with 1% paraformaldhyde. The samples were analyzed on a FACScan instrument which was standardized daily with the AutoComp software program in accordance with the manufacturer's instructions. After the instrument was set up and daily settings were recorded, each sample, stained with the CD4/CD8 reagents, was evaluated to obtain appropriate compensation settings on the instrument. The sample was then analyzed with the Simulset software after the CD45/CD14 gating reagent was run. At a minimum, the leukocyte gate purity was established as 90% and lymphocyte recovery was 95%. If these gate settings were not met by the automated program, operator-defined gates were utilized. After the appropriate analysis gates were set, cursor settings were established for fluorescence measurements with the isotype control. The remaining tubes were analyzed by the automated software program. This analysis gave a standardized approach in evaluating samples at the various sites and allowed detection of either instrument or sample preparation variability with the proficiency testing samples.

Statistical methods. In accordance with the recommendation of Gelman and Eudey (13), statistical methods that are resistant to the effects of outliers were employed for most of the data analysis. Therefore, the median, rather than the mean, was used as a measure of central tendency for the overall distributions of lymphocyte subsets. An estimate of interlaboratory variation for each subset was obtained by calculating the robust coefficient of variation (RCV) across laboratories for each sample and then finding the median of the estimates. The RCV is defined as the median absolute deviation divided by the median. The median absolute deviation is 1.5 times the median of the absolute values of deviations from the median. Friedman's rank test (11) for randomized blocks was used to compare measurements among laboratories and to compare RCVs among phenotypes. The mixed linear model that is often used for analyses such as these was also considered (21). Analysis of residuals, however, indicated that the data did not fit the assumptions of the linear model, and transformations did not reduce the problem. A multistage procedure was used for pairwise comparisons of laboratories to identify differences that contributed to any statistically significant Friedman's tests. This procedure controls the overall type 1 error rate for the 10 pairwise comparisons for each subset (42). Median residuals, 10th percentiles, and 90th percentiles were calculated for each subset in each laboratory. Tenth and 90th percentiles for the residuals were estimated directly from the observed distributions. The residuals themselves were obtained by subtracting the median of the five observations for a subset in a sample from each of the observations. The "lymphosum," which is the sum of the percentages of CD3, CD19, and CD16/56, is expected to be close to 100, so residuals were formed by subtracting 100 from the sum (36).

In accordance with the recommendation of Bland and Altman (3), the relationship between CD3$^+$ CD8$^+$ cells and total CD8 cells was examined graphically by plotting the difference between the two for each sample in each laboratory against the corresponding average. A linear model was employed to describe the relationship between the two subsets. The predictors in the model included total CD8, a binary indicator of HIV infection status, and the interaction between total CD8 and infection status. The interaction provides a test for a difference between slopes of CD3$^+$ CD8$^+$ and total CD8 in samples from HIV-infected and uninfected individuals. The binary indicator of infection status provides a test for a difference between the intercepts. Interlaboratory variation in CD3$^+$ CD8$^+$ was compared with variation in total CD8 through a Wilcoxon rank sum test of differences between the RCVs of the two subsets.

The SAS software package was used for all data analysis. The REGWF option was employed for the paired comparisons that followed Friedman's tests. All of the statistical tests used are two sided.

RESULTS

Variability in CD8$^+$ lymphocyte subset analysis. The period of analysis for the samples obtained from the proficiency testing program was September 1991 through December 1993. Over this period, the total number of samples sent for evaluation was 84, consisting of whole-blood samples obtained from 44 HIV-positive and 40 HIV-negative donors. Samples which represented duplicates or triplicates obtained from the same donor were excluded from this analysis. There were a total of 3 months in which triplicate samples were evaluated and 1 month in which duplicate samples were evaluated, giving a maximum number of both HIV-positive and HIV-negative samples of 73. This represents a total of 365 samples analyzed at five laboratories. The maximum sample size was not realized for any of the subsets being evaluated, because of either problems in sample shipment or unavailability of reagents. The specific sample size for each lymphocyte subset evaluated is shown in Table 1 and includes only samples for which we received results from all five laboratories. In the event that any one of the five laboratories did not produce a value for the individual marker for a given sample, that sample was not included in the analysis. Indeed, laboratory 5 appeared to have more missing values than the remaining sites and contributed most times to the reduced number of samples available for valid analysis. The median percentages from the five laborato-

TABLE 1. Sample sizes and median lymphocyte percentages

HIV infection status	No. of samples (median lymphocyte %)									
	CD3 CD4	CD8 CD57	CD8 DR	CD3 CD8	CD3	CD19	CD16 CD56	CD8 CD38	CD4	CD8
+	35 (21)	31 (27)	33 (37)	35 (61)	35 (84)	35 (7)	33 (7)	35 (47)	32	32
−	28 (52)	26 (7)	25 (7)	28 (24)	28 (78)	28 (12)	27 (9)	28 (18)	26	26

ries for the various subsets are shown in Table 1, demonstrating the overall trend for lower $CD3^+$ $CD4^+$ T-cell values in the HIV-positive samples with elevated total $CD3^+$ $CD8^+$ T-cell numbers and increased levels of activated CD8 subsets as shown by the $CD8^+$/HLA-DR^+ and $CD8^+$/$CD38^+$ subsets. In addition, there was an increased percentage of $CD57^+$/$CD8^+$ cells among the HIV-positive donors. The median values for total $CD3^+$ T cells in the HIV-positive samples tended to be higher than the values for the HIV-negative samples, while the B-cell and NK cell values were similar. Median RCVs were calculated as measures of the variability of these subsets across the five WITS sites. RCVs were determined for individual subsets (Table 2). A comparison of the median counts for HIV-positive and HIV-negative samples (Table 1) with the corresponding median RCVs (Table 2) shows that the median RCV varies inversely with the median percentage for all lymphocyte subsets in this analysis. Although not shown, plots of the RCV against the median for individual samples show that the RCV is inversely related to the median when the median is less than 20. Above this value, there is little correlation between the RCV and the median. The inverse relationship at low relative counts reflects the rounding error that is inherent in reporting the results to the nearest whole percentage. Rounding error elevates the RCV at low percentages even if the true RCV is constant across the range of data. Statistically significant differences in RCVs among the subsets were identified for both HIV-positive and HIV-negative samples (Friedman's test; HIV positive, $P < 0.001$; HIV negative, $P < 0.001$). Pairwise comparisons among lymphocyte subsets showed that the RCVs are, on average, higher for B cells, NK cells, and the CD8 subsets for CD57, HLA-DR, or CD38 than for $CD3^+$ $CD4^+$ or $CD3^+$ $CD8^+$.

The interlaboratory variation described above could arise from random variation within laboratories, systematic differences among laboratories, or a combination of the two. Friedman's test was employed to test for systematic differences among the laboratories. Statistically significant differences among laboratories were detected in 19 of 20 (represents a total of 10 tubes for both HIV-positive and HIV-negative subjects) tests when the test was applied separately to samples from HIV-infected ($n = 10$ tests) and uninfected ($n = 10$ tests) volunteers ($CD3^+$ $CD8^+$ in HIV-positive samples, $P = 0.0416$; $CD4^+$ in HIV-positive samples, $P = 0.0113$; $CD3^+$ in HIV-negative samples, $P = 0.0223$; $CD4^+$ in HIV-negative samples, $P = 0.0344$; all others, $P < 0.001$). The lone exception was total $CD8^+$ cells in samples from HIV-positive individuals ($P = 0.1610$). The results indicate that the systematic component of interlaboratory variation is large enough to be detected at these sample sizes.

TABLE 2. Median RCVs

Subset	Median RCV		
	HIV-infected samples	Uninfected samples	All samples
CD3 CD4	7.14	3.00	5.36
CD8 CD57	15.50	14.30	15.00
CD8 DR	16.30	33.30	20.90
CD3 CD8	4.69	5.36	5.00
CD3	3.45	1.95	1.97
CD19	12.50	12.00	12.50
CD16 CD56	21.40	11.50	14.30
CD8 CD38	7.32	14.60	10.50
CD4	6.53	3.30	5.00
CD8	4.03	4.69	4.51
Lymphosum	2.94	1.53	1.53

The differences that were detected could reflect departure of a single laboratory from the other four. This can happen only if no more than 4 of the 10 possible pairwise comparisons among laboratories for a subset are statistically significant. In 6 of 19 cases, at least five differences among laboratories were identified ($CD3^+$ $CD4^+$ in HIV-positive samples, $CD19^+$ in HIV-positive samples, and $CD8^+$ $CD57^+$ and $CD8^+$ HLA-DR^+ in both). In one case ($CD3^+$ $CD8^+$ in HIV-positive samples), only one statistically significant pairwise comparison was identified. Among the 12 subsets for which two, three, or four paired comparisons were statistically significant, there were seven cases in which a single laboratory differed from one or more of the others ($CD16^+$ $CD56^+$ in HIV-positive samples, CD4 in HIV-negative samples, CD19 in HIV-negative samples, and CD3 and $CD8^+$ $CD38^+$ in both). However, the identity of the laboratory varied among subsets. Only one of the five WITS laboratories was not an outlier in these comparisons.

While the results indicate multiple systematic differences among the laboratories, the differences for most subsets are, on average, rather small. To illustrate this point, median residuals are provided in Tables 3 and 4, along with 10th and 90th percentiles. Most of the medians are zero or close to zero. Only 10 of the 100 medians have values of at least 2%. Five of these are attributable to laboratory 5. Comparison of 10th and 90th percentiles indicates little variation in dispersion among laboratories. The exceptions, again, tend to be clustered in laboratory 5.

$CD3^+$ $CD8^+$ cells versus total CD8 ($CD3^+$ $CD8^+$ plus $CD3^-$ $CD8^+$) cells. A plot of the difference between $CD3^+$ $CD8^+$ and $CD8^+$ cells against the mean of the two shows that 301 (95.6%) of 315 point differences fall within 2 standard deviations of the mean difference (Fig. 1). This is very close to the expected percentage for normally distributed data.

The scatter in this plot indicates that the difference between $CD3^+$ $CD8^+$ and total CD8 cells is correlated with the mean. This could indicate that the relationship between $CD3^+$ $CD8^+$ and total CD8 cells varies systematically with total CD8 cells. The linear regressions that were described earlier were employed to investigate this possibility. Parameter estimates for the regressions of $CD3^+$ $CD8^+$ cells on total CD8 cells are provided in Table 5, and regression lines are plotted in Fig. 2. A statistically significant difference between the slopes of the regressions for HIV-positive and HIV-negative samples was

TABLE 3. Median residuals and 10th and 90th percentiles for uninfected subjects

Lymphocyte subset	10th percentile for residuals at site:					Median for residuals at site:					90th percentile for residuals at site:				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
CD3 CD4	−1	−1	−5	−2	−3	0.0	1.0	−0.5	0.0	0.0	4	4	1	3	2
CD8 CD57	−2	−3	−1	−1	0	−1.1	−1.0	0.0	0.0	1.5	0	0	1	2	4
CD8 DR	−5	−3	−3	−1	0	−3	0	0	1	5	0	0	2	4	11
CD3 CD8	−3	−2	−1	−1	−1	−0.5	0.0	0.0	0.0	1.0	1	2	2	2	4
CD3	−2	−4	−6	−2	−3	0.0	0.0	−1.0	0	0.0	4	2	1	5	3
CD1	−1	−2	−1	−5	−4	0.0	0.0	0.5	−1.5	0.0	2	2	2	1	2
CD1656	−2	−3	0	−2	−2	−1	0	1	0	0	0	1	3	1	2
CD8 CD38	−4	−4	−3	−3	0	−1.0	0.0	0.0	0.0	4.5	2	3	5	4	10
CD4	−1	−1	−6	−5	−5	0.0	0.0	0.0	0.0	0.0	2	4	2	3	3
CD8	−4	−1	−2	−3	−2	0.0	−1	0.0	0.0	1	1	0	2	2	4

detected (*t* test, $P = 0.0015$). The slopes in both groups differed from 1.0 (*t* test, $P < 0.0001$ for each). Thus, the proportion of total CD8 cells that were $CD3^+$ varied with infection status and also varied systematically within each infection status.

Wilcoxon's signed ranked test was employed to compare interlaboratory variation of $CD3^+$ $CD8^+$ cells with that of total CD8 cells. Comparisons of RCVs provided little evidence of systematic differences in interlaboratory variation between $CD3^+$ $CD8^+$ and total CD8 cells (HIV-positive samples, $P = 0.23$; HIV-negative samples, $P = 0.78$). However, comparisons based on the ordinary coefficient of variation (CV) led to a different conclusion. CVs for samples from uninfected individuals were, on average, greater for $CD3^+$ $CD8^+$ cells than for total CD8 cells ($P = 0.0182$). Although it is not statistically significant, the CVs for samples from HIV-infected individuals show a trend in the same direction ($P = 0.0608$). A comparison of RCVs for the combined data failed to detect a difference ($P = 0.28$), but a test based on the CVs did indicate that measurements of $CD3^+$ $CD8^+$ cells were more variable than measurements of total CD8 cells ($P = 0.0025$). The different results can be attributed to outliers that tend to inflate the CV but not the RCV. This illustrates the utility of the RCV for measurement of interlaboratory variation.

The lymphosum. The lymphosum is defined as the sum of the percentages of CD3, CD19, and CD16/56 cells (36). These subsets constitute the three major groups of lymphocytes, so the sum should be close to 100. Variation in the lymphosum could, therefore, provide a simple device for quality control.

The median lymphosum across all laboratories was 99 for the samples from HIV-infected individuals and those from uninfected individuals. Tenth and 90th percentiles indicated that, overall, the range of variation was rather narrow (10th to 90th percentiles: HIV-positive samples, 95 to 104; HIV-negative samples, 95 to 103). This narrow range of variation resulted in small median RCVs (HIV-positive samples, 2.94; HIV-negative samples, 1.53). Both medians were smaller than the corresponding values for the three subsets that make up the lymphosum. Friedman's test failed to detect any statistically significant variation among laboratories in the lymphosums (HIV-positive samples, $P = 0.18$; HIV-negative samples, $P = 0.52$). However, we did find that the laboratory with the largest residuals for the CD8 subsets (HLA-DR and CD38) did have the widest range for the lymphosum in the HIV-positive samples.

DISCUSSION

The major goal of this study was to determine the variability in the measurement of CD8 cell subsets, including activation markers (HLA-DR and CD38) and the $CD8^+$ $CD57^+$ subset, which has been associated with different functional activities (35) in HIV-infected individuals. The evaluation of the entire sample population for this study showed RCVs for these three phenotypic markers ranging from 10.5 to 20.9. These RCVs were significantly higher than that of either the $CD3^+$ $CD4^+$ or the $CD3^+$ $CD8^+$ subset. A portion of the relatively higher RCVs seen with markers with low relative counts (i.e., CD4 in HIV-positive individuals or $CD8^+$ HLA-DR^+ in HIV-negative individuals) may be attributed to rounding error. This tended to be most evident in samples with relative counts below 20%. Estimates of $CD3^+$ $CD4^+$ and $CD3^+$ $CD8^+$ measurement

TABLE 4. Median residuals and 10th and 90th percentiles for HIV-infected subjects

Lymphocyte subset	10th percentile for residuals at site:					Median for residuals at site:					90th percentile for residuals at site:				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
CD3 CD4	−1	−1	−2	−1	−4	0.0	0.0	0.0	0.0	−1.0	2	3	0	3	2
CD8 CD57	−8	−8	0	−5	0	−1.0	−4.0	0.0	0.0	5.0	0	0	4	3	9
CD8 DR	−12	−11	−6	−11	0	−3	−2	0.0	0.0	10	0	4	8	8	24
CD3 CD8	−4	−3	−2	−3	−2	0.0	0.0	0.0	0.0	1.0	1	6	4	3	6
CD3	−3	−2	−5	−2	−7	0.0	2	−1	0.0	0.0	2	7	2	4	4
CD19	0	−3	−1	−5	−3	0.0	0.0	0.0	−1.0	0.0	2	0	2	0	2
CD1656	−1	−6	−1	−1	−1	0.0	−1	0.0	0.0	0.0	1	0	2	2	3
CD8 CD38	−6	−9	−6	−4	−5	−1.0	−1.0	0.0	0.0	6.0	1	3	4	9	15
CD4	−3	−3	−4	−2	−5	0.0	0.5	0.0	0.0	0.0	1	2	1	3	1
CD8	−2	−4	−4	−4	−3	0.0	0.0	0.0	0.0	0.0	3	4	3	4	5

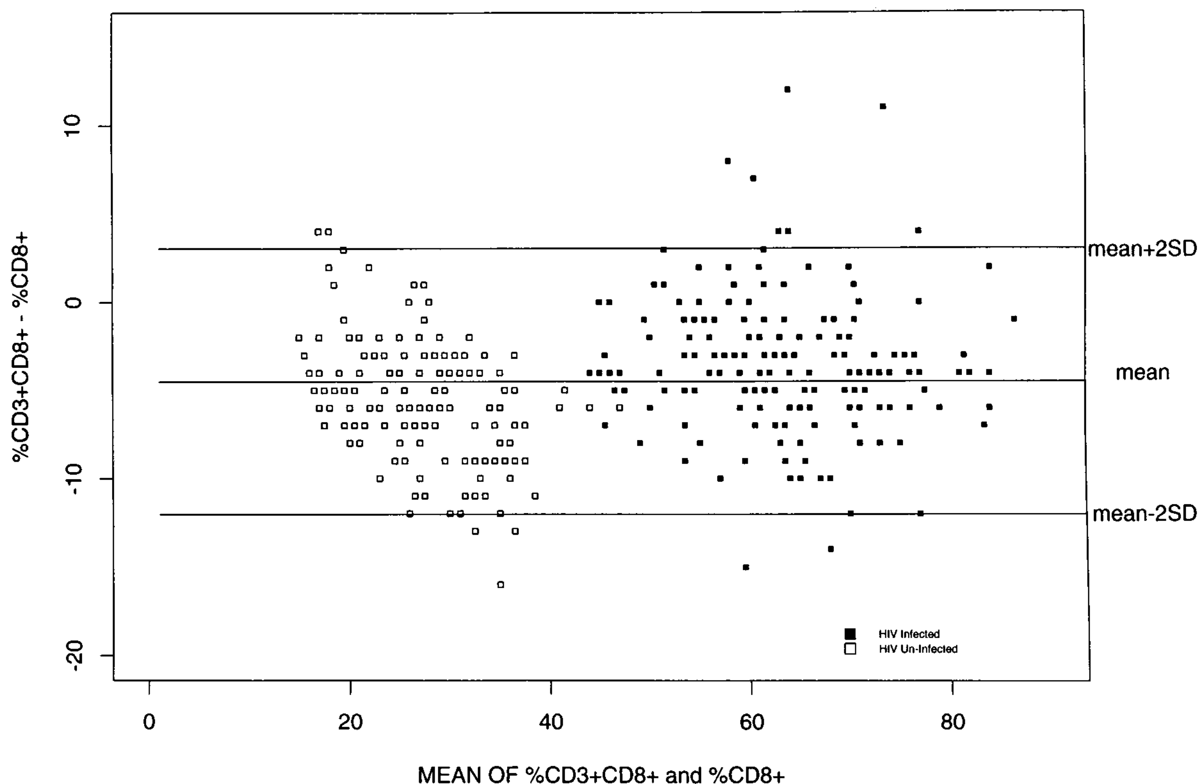

FIG. 1. Difference between percent $CD3^+$ $CD8^+$ cells and percent total $CD8^+$ cells versus the mean of the two for each sample in each laboratory. 2SD, 2 standard deviations of the difference.

variability were consistent with results of previously published reports (4, 10, 12, 14, 28, 29, 33, 34).

The major difference in analyzing these different populations is the fact that the activation and functional antigen expression of CD8 cells do not yield discretely defined populations of stained cells. This makes it more difficult in the analysis to define the cursor setting and leads to greater variability in sample analysis.

The question that arises is whether the variability in the CD8 subset analysis seen in the proficiency testing samples could be contributed by a single laboratory in this study. Evaluation of interlaboratory variability revealed that when utilizing Friedman's test, a systematic component of interlaboratory variation was detected, even with this small sample size. This did not reflect, in most cases, a difference between one laboratory and the others but was due to multiple systematic differences among the laboratories, so that biased results from a single laboratory did not influence the measurements.

Although there were statistically significant differences for the various subsets between laboratories, on average, these were very small. This point is borne out by evaluation of the median residuals shown in Tables 3 and 4, where most of the median residuals are zero. Exceptions tend to be clustered in laboratory 5.

These kinds of results demonstrate the need for caution when combining data from multiple sites. This is of significance for the WITS and other multicenter studies where there is interest in performing cross-sectional analysis of data from multiple laboratories. The fact that assays performed in a single laboratory are subject to error should not be overlooked, and the interlaboratory variability that we have measured includes this component of variation. We can only measure intralaboratory variability and its relative contribution to total variation by evaluating replicate samples. We have not analyzed sufficient numbers of replicate samples to estimate this variability. However, even if this intralaboratory error represents a substantial component of the total variability, the results of Friedman's test show that consistent differences between laboratories exist.

Another important aspect that this study has evaluated includes quality control parameters as they relate to lymphocyte subsets. A number of agencies (the Centers for Disease Control and Prevention, the National Institutes of Health, and the National Committee for Clinical Laboratory Standards) have specific recommendations for evaluation of $CD4^+$ and $CD8^+$ T cells in the monitoring of patients with HIV disease. These recommendations include utilization of the CD3 T lineage marker in conjunction with CD8 when evaluating the CD8 subset. The rationale for this is the fact that the CD8 antigen is not only expressed on T-cell subsets but is also found to be present on a proportion of NK cells. It is thought, on an empiric basis, that by including CD3 with CD8, the measure-

TABLE 5. Linear regressions of $CD3^+$ $CD8^+$ on total CD8 cells

HIV infection status	Parameter	Estimate	SE	95% Confidence limits for estimate
+	Intercept	3.6838	1.7059	0.3402, 7.0274
	Slope	0.8886	0.0258	0.8380, 0.9392
−	Intercept	1.9759	1.0942	−0.1687, 4.1205
	Slope	0.7484	0.0354	0.6790, 0.8178

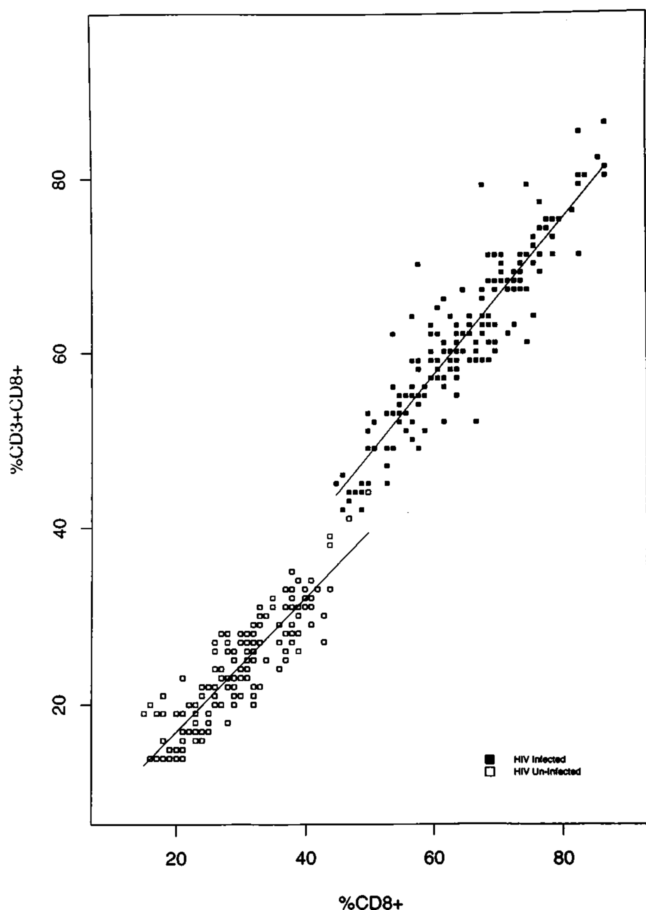

FIG. 2. Percent $CD3^+$ $CD8^+$ cells versus percent total $CD8^+$ cells, with regression lines.

ment of this marker, especially in multicenter trials, would be improved. In fact, in this study we did show a statistically significant difference between samples evaluated by $CD3^+$ $CD8^+$ or CD8 alone for both HIV-positive and HIV-negative samples. In addition, our cross-sectional analysis suggests that the proportion of $CD8^+$ cells that are $CD3^+$ does vary over the course of disease.

In summary, the data presented in this report demonstrate the need for rigorous quality control when performing immunophenotypic analysis of CD8 subsets in multiple laboratories. The importance of these types of proficiency testing procedures in the WITS comes from implementation of quality control procedures for patient results. In addition, the implications of this are important as new surrogates for monitoring of patients enrolled in HIV therapeutic trials will focus on subsets of CD8 cells, as they have been identified as important prognostic indicators in epidemiology cohorts. The Flow Cytometry Advisory Committee of the Division of AIDS has moved forward in implementing a program for certification of laboratories to include evaluation of $CD8^+$ T cells with the expectation that future studies within the AIDS Clinical Trial Group Program and other multicenter study groups will involve evaluation of CD8 subsets. These types of evaluations have already been implemented as substudies in clinical trials, and the initial results look promising with regard to the evaluation of CD8 activation markers as potential indicators of response to antiretroviral therapy. The data from the current study should help lead the way for evaluation of additional phenotypic markers on subjects enrolled in multicenter trials of antiretroviral and immunologically based therapy.

ACKNOWLEDGMENTS

This study was supported by the following contracts: N01-AI-82507 (Brigham and Women's Hospital), N01-AI-82506 (Columbia Presbyterian Hospital), N01-HD-8-2913 (State University of New York, Brooklyn), N01-AI-85005 (University of Puerto Rico), N01-AI-82505 (University of Illinois at Chicago), and N01-AI-05072 (New England Research Institute).

REFERENCES

1. **Aranda-Ansaldo, A.** 1991. A role for CD8+ T lymphocytes in the pathogenesis of AIDS. Res. Immunol. **142:**541–550.
2. **Ariyoshi, K., B. Earwood, R. Cheingsong-Popov, and J. Weber.** 1992. Is clearance of HIV-1 viremia at seroconversion mediated by neutralizing antibodies? Lancet **340:**1257–1258.
3. **Bland, J. M., and D. G. Altman.** 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet **i:**307–310.
4. **Bofill, M., G. Janossy, C. A. Lee, D. Macdonald-Burns, A. N. Phillips, C. Sabin, A. Timms, N. A. Johnson, and P. B. A. Rernoff.** 1992. Laboratory control values for CD4 and CD8 T lymphocytes. Implications for HIV-1 diagnosis. Clin. Exp. Immunol. **88:**243–252.
5. **Borrow, P., H. Lewicki, B. H. Hahn, G. M. Shaw, and M. B. A. Oldstone.** 1994. Virus-specific $CD8^+$ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J. Virol. **68:**6103–6110.
6. **Calvelli, T., T. N. Denny, H. Paxton, R. Gelman, and J. Kagan.** 1993. Guideline for flow cytometric immunophenotyping: a report from the National Institute of Allergy and Infectious Diseases, Division of AIDS. Cytometry **14:**702–715.
7. **Carmichael, A., X. Jin, P. Sissons, and L. Borysiewicz.** 1993. Quantitative analysis of the human immunodeficiency virus type 1 (HIV-1)-specific cytotoxic T lymphocyte (CTL) response at different stages of HIV-1 infection: differential CTL responses to HIV-1 and Epstein-Barr virus in late disease. J. Exp. Med. **177:**249–256.
8. **Centers for Disease Control.** 1992. Guidelines for the performance of CD4+ T-cell determinations in persons with human immunodeficiency virus infection. Morbid. Mortal. Weekly Rep. **41**(RR-8)**:**1–17.
9. **Cheingaong-Popov, R., C. Panagiotidi, B. Bowcock, A. Aronstam, J. Wadsworth, and J. Weber.** 1991. Relation between humoral responses to HIV gag and env proteins at seroconversion and clinical outcome of HIV infection. Br. Med. J. **302:**23–26.
10. **Fahey, J. L.** 1990. Doing it right: measuring T cell subsets by flow cytometry. Clin. Immunol. Immunopathol. **55:**171–172.
11. **Lehmann, E. L.** 1975. Nonparametrics: statistical methods based on ranks. Holden Day, San Francisco.
12. **Gelman, R., S.-U. Cheng, P. Kidd, M. Waxdal, and J. Kagan.** 1993. Assessment of the effects of instrumentation, monoclonal antibody and fluorochrome on flow cytometric immunophenotyping: a report based on two years of the NIAID DAIDS flow cytometry quality assessment program. Clin. Immunol. Immunopathol. **66:**150–162.
13. **Gelman, R., and L. Eudey.** 1989. Biostatistical considerations for quality assessment of immunologic measurements used in clinical and longitudinal studies. Clin. Immunol. Immunopathol. **52:**28–37.
14. **Giorgi, J. V., H.-L. Cheng, J. B. Margolick, K. D. Bauer, J. Ferbas, and N. Waxdal.** 1990. Quality control in the flow cytometric measurement of T-lymphocyte subsets: the multicenter AIDS cohort study experience. Clin. Immunol. Immunopathol. **55:**173–186.
15. **Giorgi, J. V., and R. Detels.** 1989. T-cell subset alterations in HIV-infected homosexual men: NIAID multicenter AIDS cohort study. Clin. Immunol. Immunopathol. **52:**10–18.
16. **Giorgi, J. V., H.-N. Ho, K. Hirji, C.-C. Chou, L. E. Hultin, S. O'Rourke, L. Park, J. B. Margolick, J. Ferbas, J. P. Phair, and the Multicenter AIDS Cohort Study Group.** 1994. CD8+ lymphocyte activation at human immunodeficiency virus type 1 seroconversion: development of HLA-DR+ CD38− CD8+ cells is associated with subsequent stable CD4+ cell levels. J. Infect. Dis. **170:**771–781.
17. **Ho, H.-N., L. E. Hultin, R. T. Mitsuyasu, J. L. Natud, N. A. Hausner, D. Bockstoce, C.-C. Chou, B. O'Rourke, J. M. G. Taylor, and J. V. Giorgi.** 1993. Circulating HIV-specific CD8+ cytotoxic T cells express CD38 and HLA-DR antigens. J. Immunol. **150:**3070–3079.
18. **Hoffenbach, A., P. Langlade-Demoyen, G. Dadaglio, E. Vilmer, F. Michel, C. Nayaud, B. Autran, and F. Plata.** 1989. Unusually high frequencies of HIV-specific cytotoxic T lymphocytes in humans. J. Immunol. **142:**452–462.
19. **Kestens, L., G. Vanham, P. Gigase, G. Young, I. Hannet, F. Vanlangendonck,**

F. Hulstaert, and B. A. Bach. 1992. Expression of activation antigens, HLA-DR and CD38, on CD8 lymphocytes during HIV-1 infection. AIDS **6:**793–797.

20. **Klatzmann, D., F. Barre-Sinoussi, N. T. Nugeyre, C. Dauguet, E. Vilmer, C. Griscelli, F. Brun-Vezinet, C. Rouzioux, J. C. Gluckman, J.-C. Cheman, and L. Montagnier.** 1984. Selective tropism of lymphadenopathy associated virus (LAV) for helper inducer T lymphocytes. Science **225:**59–63.
21. **Kristiansen, S.** 1991. A statistical procedure for the estimation of accuracy parameters in interlaboratory studies. Statist. Med. **10:**843–854.
22. **Landay, A. L., C. E. Mackewicz, and J. A. Levy.** 1993. An activated CD8+ T cell phenotype correlates with anti-HIV activity and asymptomatic clinical status. Clin. Immunol. Immunopathol. **69:**106–116.
23. **Levacher, N., F. Hulstaert, S. Tallet, S. Ullery, J. J. Pocidalo, and B. A. Bach.** 1992. The significance of activation markers on CD8 lymphocytes in human immunodeficiency syndrome: staging and prognostic value. Clin. Exp. Immunol. **90:**376–382.
24. **Mackewicz, C. B., H. W. Ortega, and J. A. Levy.** 1991. CD8+ cell anti-HIV activity correlates with the clinical state of the individual. J. Clin. Invest. **87:**1462–1466.
25. **Mihailov, C., A. Lamour, V. Beaudre-Bellein, N. Jezequel, M. Garre, D. Nottier, G. Guillet, and P. Youinou.** 1993. Prognostic significance of cytotoxic T cells in individuals infected with human immunodeficiency virus. J. Clin. Immunol. **13:**139–144.
26. **National Committee for Clinical Laboratory Standards.** 1992. Clinical applications of flow cytometry: quality assurance and immunophenotyping of peripheral blood lymphocytes. NCCLS tentative guideline, document H42-T, vol. 12, no. 6. National Committee for Clinical Laboratory Standards, Villanova, Pa.
27. **Pantaleo, G., S. Koenig, N. Baseler, H. C. Lane, and A. S. Fauci.** 1990. Defective clonogenic potential of CD8+ T lymphocytes in patients with AIDS. J. Immunol. **144:**1696–1704.
28. **Parker, J. W., B. Adelsberg, S. P. Azen, D. Boone, M. A. Fletcher, G. F. Gjerset, J. Hassett, J. Kaplan, J. C. Niland, T. Odom-Maryon, B. A. Operskalski, H. Prince, D. Scott, D. P. Stites, J. W. Mosley, and the Transfusion Safety Study Group.** 1990. Leukocyte immunophenotyping by flow cytometry in a multisite study: standardization, quality control, and normal values in the transfusion safety study. Clin. Immunol. Immunopathol. **55:**187–220.
29. **Paxton, H., P. Kidd, A. Landay, J. Giorgi, N. Flomenberg, E. Walker, F. Valentine, J. Fahey, and R. Gelman.** 1989. Results of the flow cytometry ACTG quality control program. Analysis and findings. Clin. Immunol. Immunopathol. **52:**68–84.
30. **Piatak, M., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K. C. Luk, B. H. Hahn, G. N. Shaw, and J. D. Lifson.** 1993. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science **259:** 1749–1754.
31. **Plata, F., B. Autran, L. P. Martins, H. S. Wain, N. Raphael, C. Mayaud, N. Denis, J. N. Guillon, and P. Debre.** 1987. AIDS virus-specific cytotoxic T lymphocytes in lung disorders. Nature (London) **328:**348–351.
32. **Prince, H. E., and E. R. Jensen.** 1991. Three-color cytofluorometric analysis of CD8 cell subsets in HIV-1 infection. J. Acquired Immune Defic. Syndr. **4:**1227–1232.
33. **Rickman, W. J., C. Monical, and M. J. Waxdal.** 1993. Improved precision in the enumeration of absolute numbers of lymphocyte phenotypes with long-term monthly proficiency testing. Ann. N. Y. Acad. Sci. **677:**53–58.
34. **Rickman, W. J., M. J. Waxdal, C. Monical, J. D. Damato, and D. S. Burke.** 1989. Department of Army lymphocyte immunophenotyping quality assurance program. Clin. Immunol. Immunopathol. **52:**85–95.
35. **Sadat-Sowti, B., P. Debre, T. Idziorek, J.-M. Guillon, F. Hadida, E. Okzenhendler, C. Katlama, C. Mayaud, and B. Autran.** 1991. A lectin-binding soluble factor released by CD8+ CD57+ lymphocytes from AIDS patients inhibits T cell cytotoxicity. Eur. J. Immunol. **21:**737–741.
36. **Schenker, E. L., L. E. Hultin, R. D. Bauer, J. Ferbas, J. B. Margolick, and J. V. Giorgi.** 1993. Evaluation of a dual-color flow cytometry immunophenotyping panel in a multicenter quality assurance program. Cytometry **14:** 307–317.
37. **Vanham, G., L. Kestens, G. Penne, C. Goilav, P. Gigase, R. Colebunders, N. Vandenbruaene, J. Goeman, G. Van Der Groen, and J. L. Ceuppens.** 1991. Subset markers of CD8 (+) cells and their relation to enhanced cytotoxic T-cell activity during human immunodeficiency virus infection. J. Clin. Immunol. **11:**345–356.
38. **Walker, B. D., S. Chakrabarti, B. Moss, T. J. Paradis, T. Flynn, A. G. Durno, R. S. Blumberg, J. C. Kaplan, M. S. Hirsch, and R. T. Schooley.** 1987. HIV-specific cytotoxic T lymphocytes in seropositive individuals. Nature (London) **328:**345–348.
39. **Walker, C. M., A. L. Erickson, F. C. Hsueh, and J. A. Levy.** 1991. Inhibition of human immunodeficiency virus replication in acutely infected CD4+ cells by CD8+ cells involves a noncytotoxic mechanism. J. Virol. **65:**5921–5927.
40. **Walker, C. M., D. J. Moody, D. P. Stites, and J. A. Levy.** 1986. CD8+ lymphocytes can control HIV infection *in vitro* by suppressing virus replication. Science **234:**1563–1566.
41. **Walker, C. M., D. J. Moody, D. B. Stites, and J. A. Levy.** 1989. CD8+ T lymphocyte control of HIV replication in cultured CD4+ cells varies among infected individuals. Cell. Immunol. **119:**470–475.
42. **Welsch, R. E.** 1977. Stepwise multiple comparison procedures. J. Am. Stat. Assoc. **72:**359.