ORIGINAL ARTICLE

Inter- and intra-laboratory variability of CD4 cell counts in Swaziland

Ganizani Mlawanda, *MB ChB, MSc Clinical Epidemiology, Dip HIV Man, DTM&H*
School of Health Systems and Public Health, Faculty of Health Sciences, University of Pretoria, and Royal Swaziland Sugar Corporation Medical Services Hospitals, Mhlume, Swaziland

Paul Rheeder, *MB ChB, MMed, PhD*
Jacqui Miot, *BPharm, PhD*
School of Health Systems and Public Health, Faculty of Health Sciences, University of Pretoria

Corresponding author: *G Mlawanda (mlawandag@yahoo.co.uk)*

Background. Analytical variability in CD4 enumeration is well known, but few studies from southern Africa have quantified the inter- and intra-laboratory variability in CD4 count measurements. In addition, the possible impact of time lapse after sample collection on CD4 reliability is not well understood.

Methods. A cross-sectional study was conducted at Royal Swaziland Sugar Corporation Hospital and three laboratories, Lab A (comparator), Lab B (national reference) and Lab C (rural hospital). Blood from HIV-infected individuals was collected using routine venepuncture into separate specimens for each of the three laboratories. The samples were further subdivided at each laboratory: one was run at 12 hours and the second at 24 hours after venepuncture. The results of absolute CD4 count and CD4 percentage testing were compared within (intra-laboratory) and between (inter-laboratory) laboratories.

Results. Among 53 participants, the mean CD4 count at 12 hours was 373 cells/µl, 396 cells/µl and 439 cells/µl, and at 24 hours 359 cells/µl, 389 cells/ µl and 431 cells/µl, for laboratories A, B and C, respectively. The coefficient of intra-laboratory variation was 4%, 8% and 20% for CD4 count for laboratories A, B and C, respectively. Comparing 12- and 24-hour measurements, the mean difference (bias) within the laboratories between the two time points (and limits of agreement, LOAs) was 14 (-46 to 73), 8 (-161 to 177) and 7 (20 to 33) cells/µl for labs A, B and C, respectively. Comparing Lab A versus Lab B, lab A versus Lab C and Lab B versus Lab C, the inter-laboratory bias for the CD4 count at 12 hours was -32, -64 and -38 cells/µl, respectively. The corresponding LOAs were -213 to 150, -183 to 55, and -300 to 224, respectively. At 24 hours, the biases and LOAs were similar to those at 12 hours.

Conclusions. CD4 counts appeared reliable at all three laboratories. Lab B and Lab C were clinically interchangeable with the comparator laboratory, Lab A, but not between themselves. Time to measurement does not affect the inter-laboratory agreement within 12 and 24 hours.

S Afr J HIV Med 2012;13(2):59-63.

Thirty-five million people are infected by HIV globally, two-thirds of whom live in sub-Saharan Africa.[1] Antiretroviral therapy (ART) is a critical intervention for reducing HIV-related morbidity and mortality, but delivery of ART requires multiple laboratory investigations.[2] In particular, determination of eligibility for ART initiation relies heavily on CD4 enumeration, and CD4 results are monitored as the major indicator of response to treatment over time.

The gold standard technique for CD4 enumeration is flow cytometry.[3,4] Biological and analytical (laboratory) variations are known to affect CD4 enumeration; biological factors can that influence CD4 results include haemodilution in pregnancy, seasonal and diurnal variations (lowest at approximately 12:30 pm, highest at 8:30 pm), surgery, viral infections, tuberculosis, some intercurrent illnesses, corticosteroids, interferon and cancer chemotherapy.[3]

Laboratory variations are known to occur when enumeration techniques different from the gold standard, flow cytometry, are used.[3,4] In addition, variations are known to be subject to inter-observer differences as well as inter-laboratory differences.[5] The time to performing CD4 may also cause variation in final CD4 count; the World Health Organization (WHO) therefore recommends that all CD4 counts be done within 72 hours from the time of blood collection.[3,4] In Swaziland and many other parts of southern Africa, blood for CD4 testing is collected from various health centres and then sent to central laboratories where analysis is done. The time of arrival of samples differs greatly according to distance from the laboratory, but the impact of time differences on CD4 results is not well understood.

Clinicians rely on accurate CD4 values, despite this variability, to make decisions regarding ART initiation and management. Some previous studies of CD4 variability have produced worrying results. Sax and Boswell analysed the implication of between-laboratory variations and found that 58% of CD4 count results had enough variation to have led to conflicting treatment recommendations.[6] Pattanapanyasat and Chimma found CD4 variation between CD4 cell count results conducted using flow cytometers of different ages in service.[7] Various new CD4 enumeration techniques, for example the Guava Easy CD4 and capillary-based CD4, have been compared with gold-standard techniques and found to be comparable.[3,4,8]

Ensuring accurate CD4 counts has become more important recently, since ART is being initiated at higher CD4 counts, when clinical signs tend to be less sensitive in detecting immune suppression.[2] In Swaziland, there has been widespread suspicion among HIV clinicians regarding discrepancies in CD4 count results within and between laboratories, and concern that these discrepancies may potentially be large enough to affect decisions to start ART. In order to address this problem, this study sought to evaluate the intra- and inter-laboratory variability in CD4 cell enumeration.

Methods

This study was undertaken at HIV clinics at the Royal Swaziland Sugar Corporation Hospital in Swaziland and three laboratories, Lab A, Lab B and Lab C (identity of the laboratories deliberately not disclosed). Lab A was a reputable, internationally accredited South African laboratory commonly used as standard in clinical practice across southern Africa. Lab B was the Swazi national reference laboratory based in the capital city, 250 km away from the study setting, and had a turnover of 4 000 CD4 enumerations per

Table 1. Demographic, clinical and laboratory data of participants

Demographic characteristics				
Gender (*n*)				
Male	28			
Female	25			
Age (years) (mean (SD))	37.4 (9.5)			
Weight (kg) (mean (SD))	64.8 (12.2)			
Clinical characteristics				
WHO stage (%)				
I	32.1			
II	22.6			
III	13.2			
IV	32.1			
On TB treatment (%)	11.3			
On ART (%)	47.2			
Inpatients (%)	9.4			
Outpatients (%)	90.6			
Laboratory parameters				
Lab A (52 observations)	Mean	25th centile	50th centile	75th centile
CD4 count at 12 h (cells/µl)	373	181	336	539
CD4 count at 24 h (cells/µl)	359	177	323	518
CD4 % at 12 h	17	10	15	22
CD4 % at 24 h	17	10	15	21
Lab B (52 observations)	Mean	25th centile	50th centile	75th centile
CD4 count at 12 h (cells/µl)	396	185	359	568
CD4 count at 24 h (cells/µl)	389	183	346	535
CD4 % at 12 h	18	11	17	24
CD4 % at 24 h	18	10	17	23
Lab C (51 observations)	Mean	25th centile	50th centile	75th centile
CD4 count at 12 h (cells/µl)	439	249	397	611
CD4 count at 24 h (cells/µl)	431	233	396	594
CD4 % at 12 h	18	10	16	22
CD4 % at 24 h	18	10	16	22
Mean time to running CD4 tests (h)	First CD4	Second CD4		
Lab A	12.0	24.0		
Lab B	12.0	24.0		
Lab C	12.0	25.1		

month. Lab C was a rural mission hospital laboratory located 80 km from the study site and had a turnover of 1 700 samples per month. All the three laboratories used a flow cytometric CD4 enumeration method, and trained laboratory technicians performed the CD4 tests.

To be eligible, patients had to be adults (>18 years), give informed consent to the study, and be visiting the health facility for routine CD4 count. The study included patients regardless of whether they were on ART or not. After participants' consent had been obtained, blood was collected into EDTA tubes, using routine venepuncture technique, in three aliquots, one each for Lab A, Lab C and Lab B. The samples were further split into two aliquots at each respective laboratory, one of which was run at 12 hours and the second at 24 hours after venepuncture. A reliable transport vehicle ensured that specimens reached all laboratories within stipulated time.

A sample size of 53 was used. For this type of study, Altman and Bland recommend a sample size of 30 as 'minimum acceptable' and 50 as 'good' as it gives a 95% confidence interval (CI) about ±0.34 *s*, where *s* is the standard deviation (SD) of the differences between measurements by the two methods.[9]

Data were analysed using STATA version 10. For intra-laboratory variability, the coefficient of variation (CV) and Bland-Altman (BA) method were used. The BA method was the predominant technique for inter-laboratory variability. Bland-Altman plots were generated in Excel Analyze-it. In both cases, for repeatability and agreement, comparison was based on clinically significant reference ranges used previously in most studies: 0 - 10% for CV,

Table 2. Intra-laboratory bias and limits of agreement for CD4 count and CD4 percentage at 12 and 24 hours

		Limits of agreement		Interpretation
	Bias (95% CI)	Lower (95% CI)	Upper (95% CI)	Clinically repeatable?
Absolute CD4 count				
Lab A	13.5 (5.0 to 21.9)	-46.0 (-60.6 to -31.5)	73.0 (58.5 to 87.6)	Yes
Lab C	8.2 (-16.0 to 32.4)	-160.5 (-202.2 to -118.9)	176.9 (135.3 to 218.6)	Yes
Lab B	7.0 (3.2 to 10.7)	-19.5 (-25.9 to -13.0)	33.4 (26.9 to 39.9)	Yes
CD4 %				
Lab A	0.1 (-0.2 to 0.3)	-1.7 (-2.2 to -1.3)	1.9 (1.5 to 2.4)	Yes
Lab C	-0.3 (-0.7 to 0.1)	-2.9 (-3.5 to -2.2)	2.3 (1.7 to 3.0)	Yes
Lab B	0.1 (-0.3 to 0.5)	-2.8 (-3.5 to -2.1)	3.0 (2.3 to 3.7)	Yes

*Interpretation based on comparison of limits of agreement with clinically significant range of CV <10%, and ranges for clinical significance: ±19.5% for CD4% ±250 cells/µl for CD4 count.[7,8,10,11]

Table 3. Inter-laboratory bias and limits of agreement for CD4 count and CD4 percentage at 12 and 24 hours

		Limits of agreement		Interpretation*
Laboratories	Bias (95% CI)	Lower (95% CI)	Upper (95% CI)	Clinically interchangeable?
CD4 count at 12 h				
Lab A/Lab B	-31.5 (-57.6 to -5.5)	-213.3 (-258.2 to -168.4)	150.2(105.3 to 195.1)	Yes
Lab A/Lab C	-64.3 (-81.6 to -47.0)	-183.8 (-213.6 to -154.0)	55.2 (25.4 to 85.0)	Yes
Lab B/Lab C	-38.2 (-75.6 to -0.6)	-300.2 (-364.8 to -235.5)	223.9 (159.2 to 288.5)	No
CD4 % at 12 h				
Lab A/Lab B	-1.2 (-2.7 to 0.3)	-11.7 (-14.3 to -9.1)	9.3 (6.7 to 11.9)	Yes
Lab A/Lab C	-0.7 (-1.1 to -0.4)	-3.1 (-3.7 to -2.5)	1.7 (1.1 to 2.2)	Yes
Lab B/Lab C	0.5 (-1.1 to 2.1)	-10.7 (-13.4 to -7.9)	11.6 (8.9 to 14.4))	Yes
CD4 count at 24 h				
Lab A/Lab B	-35.6 (-60.0 to -11.1)	-205.7 (-247.6 to -163.7)	134.5 (92.5 to 176.5)	Yes
Lab A/Lab C	8.2 (-16.0 to 32.4)	-195.0 (-227.6 to -162.5)	65.8 (33.3 to 98.3)	Yes
Lab B/Lab C	7.0 (3.2 to 10.7)	-265.0 (-321.3 to -208.7)	191.4 (135.0 to 247.7)	No
CD4 % at 24 h				
Lab A/Lab B	-1.2 (-2.5 to 0.2)	-10.5 (-12.8 to -8.2)	9.2 (5.7 to 10.5)	Yes
Lab A/Lab C	-1.1 (-1.6 to -0.5)	-4.9 (-5.8 to -3.9)	2.7 (1.8 to 3.6)	Yes
Lab B/Lab C	0.1 (-1.3 to 1.5)	-9.7 (-12.1 to -7.3)	9.9 (7.4 to 12.3)	Yes

*Interpretation based on comparison of limits of agreement with clinically significant range of CV <10%, and ranges for clinical significance: ±19.5% for CD4 % and ±250 cells/µl for CD4 count.[7,8,10,11]

±250 cells/µl for CD4 count and 19.5% for CD4 percentage.[7,8,10,11] Clinical impact on antiretroviral therapy (ART) initiation was assessed by Kappa coefficients with comparison to the standard reference scales.[12]

Results

Fifty-three participants consented to participate in the study. The mean CD4 count was 373 cells/µl, 396 cells/µl and 439 cells/µl at 12 hours, and 359 cells/µl, 389 cells/µl and 431 cells/µl at 24 hours, for Lab A, Lab B and Lab C, respectively. Subsequent Wilcoxon sign-rank test revealed some statistically significant differences in CD4 count between the laboratories. Table 1 summarises the demographic, clinical and laboratory characteristics of participants.

Intra-laboratory variability. The CV for CD4 count for Lab B was low (3.4%) compared with Lab A (8.5%). This was consistent with intra-laboratory repeatability based on clinically significant CV range of 0 - 10%. For Lab C the CV was 20.1%, a finding consistent with poor repeatability. For all three laboratories, the CV of CD4 percentage was even lower: 5.6%, 8.34% and 7.5% for Lab A, Lab B and Lab C, respectively. The results using the BA method showed that both CD4 count and CD4 percentage were repeatable, when compared with clinically significant ranges ±250 cells/µl and ±19.5%, for all the laboratories: for CD4 count, the limits of agreement were -46 cells/ µl to 73 cells/µl for Lab A, -20 cells/µl to 33 cells/µl for Lab B, and -161 cells/µl to 177 cells/µl for Lab C, as per Fig. 1 and Table 2. The BA plots for Lab A, Lab B and Lab C had no dispersion suggesting evidence of systematic error.

Inter-laboratory agreement at 12 hours. For CD4 count, at 12 hours, both Lab C and Lab B could be clinically interchanged with the comparator, Lab A, based on the limits of agreement which fell within the clinically significant range (defined as ±250 cells/µl): -184 cells/ µl to 55 cells/µl for Lab C, and -213 cells/µl to 150 cells/µl for Lab B, which was much wider than for Lab C. When Lab B was compared for agreement with Lab C, the limits of agreement were -300 cells/ µl to 224 cells/µl, which were out of the clinically significant range, and we therefore concluded that the two laboratories could not be clinically interchanged. For CD4 percentage all the laboratories could be clinically interchanged. Compared with the comparator, Lab A, the limits of agreement for Lab B were -12 cells/µl to 9 cells/µl and -3 cells/µl to 2 cells/µl for Lab C; between Lab B and Lab C the limits were -11 cells/µl to 12 cells/µl. Table 3 summarises the results for inter-laboratory variability based on BA results at 12 hours and at 24 hours.

Inter-laboratory agreement at 24 hours. Time to measurement had no significant impact on inter-laboratory agreement based on the limits of agreement and biases at 24 hours were similar to those at 12 hours for both CD4 count and CD4 percentage. When compared with Lab A, the limits of agreement at 24 hours were -205 cells/µl to 135 cells/µl for Lab B and -195 cells/µl to 66 cells/µl for Lab C. For Lab B/Lab C the limits of agreement were -265 cells/µl to 191 cells/µl. For CD4 percentage, all the laboratories were clinically interchangeable. The limits of agreement were -11% to 9% for Lab A/Lab B, -5% to 3% for Lab A/Lab C and -10% to 10% for Lab B/Lab C, which were within the reference range, ±19.5%.

Clinical impact on ART initiation. Compared with Lab A, the percentage agreement for ART eligibility was 81% (i.e. 19% of patients were misclassified) for Lab B and 89% (11% of patients misclassified) for Lab C. For Lab A/Lab B, 23% eligible patients would be misclassified and not initiated on ART, as shown in Table 4.

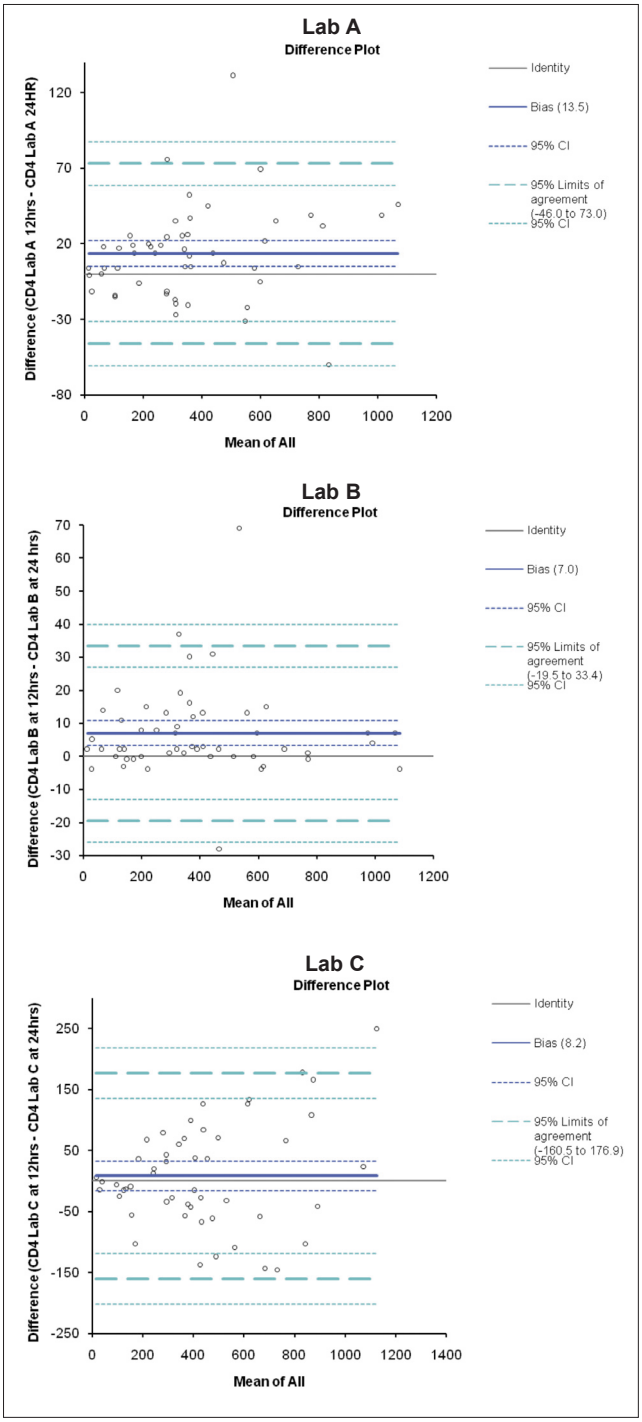

Fig. 1. Bland Altman plots for intra-laboratory variability of CD4 count for Lab A, Lab B and Lab C.

Discussion

In this study we looked at intra- and inter-laboratory variability, a topic that has been investigated previously but for which there are few data from southern Africa.[5-8,11,13] We also analysed the impact of time to measurement on the eventual CD4 result, both within the same laboratory and across participating laboratories. CD4 count had good repeatability for all the three laboratories, based on preset clinically significant ranges. Likewise, CD4 percentage had minimal variation for

all the laboratories and even lower CV, a sign of stronger repeatability than for CD4 count. These findings concurred with previous intra-laboratory studies.[7,8,10,11]

Inter-laboratory variability. Several studies on inter-laboratory and inter-method variability of CD4 count have been published and most show good agreement and interchangeability.[7,10,11] Two studies, however, found significant variations across different laboratories.[5,13] In this study, inter-laboratory clinical interchangeability results at 12 and 24 hours showed that agreement was independent of time to measurement. The limits of agreement were similar when time to measurement was 12 hours or 24 hours. This finding mirrors the WHO laboratory recommendation that CD4 remains stable within 72 hours from time of venepuncture.[3,4] Clinicians using the laboratories in this study should therefore trust equally CD4 results done at 12 hours and 24 hours.

For CD4 percentage, both Lab B and Lab C were in agreement with the comparator laboratory, Lab A, at 12 and 24 hours with narrower limits of agreement than for CD4 count. Once again, stability of CD4 percentage and agreement with the comparator laboratory make it a potentially trustworthy and stable parameter to use in our setting for possible inclusion in guidelines to determine when to start ART, as suggested in some previous studies.[8,10]

The degrees of misclassification in this study were similar to findings from a study by Thakar and Kumar, which found a kappa factor range of 74% for a CD4 count below 350 cells/µl when two laboratories were being compared.[11] Repeating CD4 count measurement and not relying on single CD4 count results have been known to reduce disease misclassification.[6] One shortfall of this use of misclassification as done here is that it does not differentiate between low-magnitude inaccuracy, for example a count of 349 cells/µl being misclassified as >350 cells/µl, which may be reasonably expected from any test, and high-magnitude inaccuracy. A study that includes many CD4 values falling close to the defined cut-off (as measured by the reference test) will show higher rates of misclassification by the new test than a study in which the majority of values lie away from the threshold.[4]

The clinically significant ranges used in this study were ±250 cells/µl, ±19.5% and CV <10%, because these were the ranges used in similar studies which had pre-defined ranges.[7,8,10,11] The results of repeatability and agreement therefore relied on this pre-defined range. However, the choice of clinically significant ranges is debatable, and a narrower range of ±100 cells/µl could have changed the interpretation of these results greatly. However, the magnitude of CD4 count or CD4 percentage variability that can affect clinical decision making remains poorly defined.[13] The author felt that based on the new ART initiation threshold, 350 cells/µl, a range of ±250 cells is reasonable.

In conclusion, CD4 count and CD4 percentage appeared to be repeatable for all the three laboratories. Lab B and Lab C were clinically interchangeable with the comparator laboratory, Lab A, for both CD4 count and CD4 percentage but not between themselves. Time to measurement does not affect the inter-laboratory agreement within 12 and 24 hours. The clinical implications of inter-laboratory variation on disease misclassification were comparable to those from previous studies.

Table 4. Impact of CD4 variations at ART initiation threshold on treatment decision

Laboratories	Agreement* (%)	Expected agreement (%)	Kappa	Misclassified (%)
Lab A/Lab B	81.1	48.4	0.6	18.9
Lab A/Lab C	88.7	49.9	0.8	11.3
Lab B/Lab C	77.4	50.2	0.6	22.6

*Strength of agreement according to Byrt's criteria for assessing Kappa strength: excellent agreement = 0.93 to 1; very good agreement = 0.81 to 0.92; good agreement = 0.61 to 0.80; fair agreement = 0.41 to 0.60; slight agreement = 0.21 to 0.40; poor agreement = 0.01 to 0.20; no agreement <0.00.[12]

Sources of support and disclosure of funding. A University of Pretoria RESCOM grant partially covered the laboratory costs during the study.

Competing interests by authors. None.

Research ethics committee approval. Ethics approval was given by the University of Pretoria Ethics Committee 74/2010 on 21 April 2010, and institutional ethical approval was also obtained.

Author contributions. Dr Ganizani Mlawanda conceived the study, and formulated the study design, data collection, statistical analysis and manuscript design. Prof. Paul Rheeder and Dr Jacqui Miot were active supervisors throughout from conception to final manuscript. All authors read and approved the final manuscript.

REFERENCES

1. UNAIDS. Report on the Global HIV/AIDS Epidemic 2010: Executive Summary. Geneva: UNAIDS, 2010. http://www.unaids.org/globalreport/ (accessed 2 December 2011).
2. World Health Organization. Rapid Advice – 2009: Antiretroviral Therapy for HIV Infection in Adults and Adolescents. Geneva: WHO Press, 2009. http://www.who.int/hiv/pub/arv/rapid_advice_art.pdf (accessed December 2009).
3. World Health Organization. Laboratory Guidelines for Enumerating CD4 T Lymphocytes in the Context of HIV/AIDS-2007. New Dehli: World Health Organization, 2007. http://www.searo.who.int/LinkFiles/BCT_HLM-392.pdf (accessed 1 June 2011).
4. World Health Organization. Technical Brief on CD4 Technologies – 2010. Geneva: WHO Press, 2010. http://www.who.int/hiv/topics/treatment/tech_brief_20100501_en.pdf (accessed 1 September 2011).
5. Harwell TS, Ferbas J. Are there differences between laboratories that use of fail to use the CDC's guidelines to measure CD41 and CD81 T cells? Cytometry 1995;21:256-257. [http://dx.doi.org/10.1002/cyto.990220314]
6. Sax PE, Boswell SL. Potential clinical implications of inter-laboratory variability in CD4 cell count in patients with HIV infection. Clin Infect Dis 1995;21(5):1121-1125. [PMID 8589131]
7. Pattanapanyasat K, Chimma P. CD41 T-lymphocyte enumeration with a flow-rate based method in three flow cytometers with different years in service. Cytometry Part B 2008;74B:310-318. [http://dx.doi.org/10.1002/cyto.b20425] [PMID: 18481294]
8. Josefowicz S, Elad K. Guava Technologies EasyCD4(tm) and EasyCD8(tm) Assays vs. the BD Biosciences MultiTest(tm) Assay for the enumeration of human CD4+ and CD8+ T cells – three-site evaluation. Poster session presented at the Clinical Cytometry Meeting, Long Beach, CA, 17-19 October 2004. Poster No. 32.
9. Altman DG, Bland JM. Frequently asked questions on the design and analysis of measurement studies. [Discussion list on internet.] www-users.york.ac.uk/-mb55/meas/sizemeth.htm (accessed 30 November 2009).
10. Yu LM, Easterbrook PJ. Relationship between CD4 count and CD4% in HIV infected people. Int J Epidemiol 1997;26:1367-1371. [http://dx.doi.org/10.1093/ije/26.6.1367]
11. Thakar MR, Kumar KB. Comparison of capillary based micro-flurometric assay for CD4+ T cell count estimation with dual platform Flow cytometry. AIDS Research and Therapy 2006;3:26. [http://dx.doi.org/10.1186/1742-6405-3-26] [PMID: 17042936] [PMCID: PMC1636060]
12. Byrt T. How good is that agreement? Epidemiology 1996;7(5):561. [PMID: 8862998]
13. Pollock MA, Jefferson SG. Method comparison – a different approach. Ann Clin Biochem 1992;29:556-560. [PMID: 1444169]