The effect of Participation in a regional CD4 split-sample proficiency testing on CD4 test results at health laboratories in Western Kenya

Burmen BMBChB MPH PhDs¹, Mutai K MSc App. Stats¹, Kamire V MA Project Management¹, Frodwa N BSc

Biotech MSCs App Epidemiology and Laboratory Management, Kenya Medical Research Institute, Centers for Global Health Research, Kisumu, Kenya

Abstract- A longitudinal survey was conducted at 35 field laboratories within the former Nyanza and Western Provinces of Kenya that were served by the Kenya Medical Research Institute HIV Reference regional laboratory. This was done to evaluate the effect of participation in the KEMRI CD4 QA program on accuracy of CD4 results produced by Field Laboratories within health facilities that had consistently participated External Quality Assurance program, EQA, (>than 75% of the times) in the KEMRI Regional CD4 split-sample proficiency testing 2014-2015. The sample split-test proficiency testing model entailed splitting a sample into two, testing one at the field laboratory and the other at the reference laboratory using the FACScalibur lyse no-wash immunophenotyping technique (Gold standard) method. Each field laboratory provided five samples and the corresponding field CD4 results at each submission. Each submission was done once every two months and was followed by a site visit to return results and to mentor staff or initiate corrective action at participating laboratories. The Accuracy of CD4 test results generated by field laboratories at the beginning and the end of the participation in the PT program were compared using bias, LOA and upward misclassification probability for CD4 cut-off of 350 cells/µl. A decrease in the bias [-2.5 (95% CI -22.2; 17.1) to -1.5 (95% CI -27.1; 24.2)]; LOA [-196.9; 191.8] to (-161.0; 158.1)] and upward misclassification probability for CD4 cut-off of 350 cells/µl (27% to 17 %) was observed from the beginning to end of the participation cycles. Participating in CD4 EQA programs decreased the variability of test results between the field and reference laboratories and should be done regularly and continuously.

Index Terms- Bland-Altman, Limits of agreement, Bias, CD4 EQA, Proficiency testing

I. INTRODUCTION

In 2013, of an estimated 35 million[32.3-37.2 million] people living with HIV in the world, 71% were from sub-Saharan Africa[1]. The same year, there were approximately 1.6 PLHIV in Kenya[2]. The HIV virus attacks the CD4 cells, a sub-set of T-lymphocytes which are responsible for regulating immune responses[3]. The enumeration of CD4 cells is required for staging of disease as well as treatment monitoring. The scale-up of ART has increased the demand of CD4 testing services by HIV programs[4]. With the decentralization of ART provision and consequently CD4-testing laboratories in resource-constrained countries, it has been difficult to ensure that similar quality standards that are present in reference laboratories are established at all field laboratories[5].

It is internationally acknowledged that all forms of laboratory testing need to be subjected to Quality Assurance (QA) to enable laboratories maintain high levels of accuracy and proficiency. Internal quality control (IQC) and external quality assessment (EQA) programs are used to evaluate and continuously improve analytical quality and are valuable tools in the quality improvement process of clinical laboratory services[5,6].

Quality Assurance (QA) programs ensure that participating laboratories provide consistent, high quality results, with minimal variability between results generated at field and reference laboratories[7]. There are 3 components of an effective External Quality Assessment (EQA) laboratory program. These are:- site supervision, retesting of specimens, and proficiency testing (PT). Proficiency testing (PT) is the retesting of unknown samples by an approved PT program provider who then scores the results using the pre-set grading criteria and sends the laboratory scores reflecting how accurately it performed the testing. Accreditation organizations routinely monitor their laboratories’ performance through participation in PT[5]. Routine reviews of PT reports by the laboratory staff provides an alert to areas of testing that are underperforming and also indicate subtle shifts and trends that, over time, would affect their patient results[8]. PT is one component of quality assurance that can identify problems for follow up and corrective action to ensure accurate test results[7]. The terms External Quality Assessment (EQA) and Proficiency Testing (PT) are used interchangeably[6]. Until recently, many laboratories were not enrolled in any external CD4 quality assessment (EQA) programs such as QASI (quality assessment and standardization for immunological measures relevant to HIV/AIDS) or UK NEQAS (United Kingdom National External Quality Assessment Service) which is known to have a positive impact on CD4 quality; this issue to the costs, lack of Good Laboratory Practice training and implementation, and logistical challenges of participation [9, 10].

B. Objective Statement- The Kenya Medical Research Institute (KEMRI) CD4 EQA program was incepted in 2010 to promote the accuracy of CD4 test results, as an alternative to the established PT programs that are costly[5]. It is a modified PT system that entails the use of split-sample testing and supportive supervision to promote the quality of CD4 test results generated at field laboratories[8]. We evaluated the effect of participation...
in the KEMRI CD4 QA program on validity of CD4 results produced by laboratories in Western Kenya

II. METHODS

A. Study design and setting

A longitudinal survey was conducted at health facilities that were invited to participate in the KEMRI CD4 split-test proficiency program, a regional CD4 EQA program, on a bi-monthly basis between February 2014 and February 2015 for a total of seven rounds.

B. Facility selection

All the health facilities found within the former Nyanza and Western Provinces that were currently providing CD4 testing services were invited to voluntarily participate in the KEMRI CD4 EQA program. At the time of selection i.e. facilities that were selected to participate in the CD4 split-test Proficiency testing program run by KEMRI had to be of high volume (tested at least 1000 or more CD4 samples per month for three months preceding August 2012), within 100 km radius of the Reference Laboratory in Kisumu and to have consistently participated in the KEMRI/CDC CD4 IQA (>than 75% of the times) during the period. 54 facilities in the region met these criteria. For this analysis, we excluded facilities that had participated in less than three cycles leaving 35 facilities; this was done based on published literature that showed performance of participating laboratories would only increase after participation in three rounds [11]. Logistical problems, absence of reagents, absence of key staff, and machine breakdown may have prevented full participation (KEMRI HISS, Laboratory Manager, Personal Communication, 4th July 2016). For this analysis, the facilities were not sampled, all the facilities that were participating in the provision of CD4 testing and services within the former Nyanza and Western Provinces of Kenya were included in the analysis.

C. Procedures

The sample split-test proficiency testing model entailed splitting a sample into two, testing one at the field laboratory and the other at the reference laboratory using the FACScalibur lyse no-wash immunophenotyping technique (Gold standard) method[8]. Each field laboratory provided five samples and the corresponding field CD4 results at each submission (a participating round). Each submission was followed by a site visit to return result and to mentor staff or initiate corrective action at participating labs.

D. Sample collection, processing and transport

Whole blood samples were collected in an ethylenediaminetetraacetic acid (EDTA) or Becton Dickinson (BD) CD4 stabilization tube, from persons aged 5 years and older, had an accompanying CD4 result from the each of the study laboratories and were within the tube manufacturer’s recommended period of collection (48 hours for those collected in EDTA tubes and 7 days for those collected in BD CD4 stabilization tubes). Samples were excluded or rejected if poorly labeled or clotted. This is because they could not be processed to obtain a second CD4 result; a requirement for inclusion in this analysis. All equipment at the field laboratories used automatic gating strategies to identify CD4 T cells. All flow cytometers underwent annual preventive maintenance where laser alignment and calibration and spectral compensation were done.

Samples from the study laboratories were transported at room temperature by courier or delivered in person to the reference laboratory. The time between sample collection and processing may have varied between the facility laboratories; however, as documented in the literature, inter-laboratory agreement of CD4 test results are independent of time to measurement [12]. Laboratory personnel at the reference laboratory who processed the study samples were blinded to the CD4 results from the participating field laboratories.

E. Data collection

At each participating round, all facilities had to attach a sample sending form that included the following information; the name of the laboratory, type of laboratory, the type of CD4 machine used, the number of CD4 samples received and processed during the month of submission, the number of peripheral laboratories networked to the submitting (field) laboratory, the sample ID of each submitted sample, the age and gender of the patient from whom the sample was drawn, the test lab CD4 test result as reported by the submitting laboratory and the QA lab absolute CD4 count as reported by the reference (index) laboratory. All facilities were anonymized as per protocol. Data was entered into MS Excel spreadsheets and imported into R Version 3.2.1 for analysis[13].

F. Definitions

All health facilities that participated in the program were referred to as field laboratories; field laboratories were categorized into those run by the ministry of health in Kenya (government-owned) and others (run by faith based organizations and non-governmental organizations). The monthly workload at the test laboratories was categorized as high workload (≥1000 samples per month) or low workload (<1000 samples per month). The number of peripheral laboratories networked to a central laboratory were categorized as <12 and ≥12 laboratories.

G. Data analysis

Results from field laboratories were paired with those from reference and then compared for accuracy. The lab results at the reference laboratory were used as the ‘gold standard’ based on the ISO certification of the KEMRI reference lab and the inherent characteristics of split testing which requires one result to be used as a reference [14]. We compared results for field laboratories to those of the reference laboratories at the beginning and the end of the program using different measures. Similarly, we compared laboratories who had participated in ≤4 rounds to those that had participated in >4 rounds at the end of the program.

To determine the reproducibility of CD4 test results by field laboratories, agreement between results from the field laboratories and reference laboratory was evaluated using Bland Altman analysis. A mean difference (bias) was computed for each sample submitted to measure the level of accuracy and if the confidence interval included zero then there was no evidence of systematic bias between the two measurements i.e., the field and reference laboratory measurements. The limits of agreement

www.ijsrp.org
B.Field laboratories’ participation

During the entire CD4 program, six field laboratories participated in 7 rounds, five participated in 6 rounds, four participated in 5 rounds, eight participated in 4 round and twelve participated in 3 rounds. In summary, 15 (43%) field laboratories participated in 5 or more cycles (Table 1). Data not shown indicates that bias in the results was not significantly affected by patient factors (i.e. age and gender), machine type, workload and number of networked sites.

<table>
<thead>
<tr>
<th>Facility characteristics</th>
<th>n (%) or Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facility type</td>
<td></td>
</tr>
<tr>
<td>Government facilities</td>
<td>21 (60%)</td>
</tr>
<tr>
<td>Other facilities*</td>
<td>14 (40%)</td>
</tr>
<tr>
<td>Machine type</td>
<td></td>
</tr>
<tr>
<td>BD FACSCALIBUR</td>
<td>3 (9%)</td>
</tr>
<tr>
<td>CYFLOW</td>
<td>14 (40%)</td>
</tr>
<tr>
<td>FACS COUNT</td>
<td>18 (51%)</td>
</tr>
<tr>
<td>Work load</td>
<td>400 (300-900)</td>
</tr>
<tr>
<td>Network sites</td>
<td>8 (3-21)</td>
</tr>
<tr>
<td>No of times participated</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12 (34%)</td>
</tr>
<tr>
<td>4</td>
<td>8 (23%)</td>
</tr>
<tr>
<td>5</td>
<td>4 (11%)</td>
</tr>
<tr>
<td>6</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>7</td>
<td>6 (17%)</td>
</tr>
</tbody>
</table>

*Other facilities includes those facilities run by faith based organizations (FBO), non-governmental organizations (NGO) or private individual(s)

C.Relationship between number of cycles participated in and the validity of field laboratory results at the end of the EQA program

Reproducibility of field test results

Bland-Altman analysis revealed an excellent agreement between the field and reference laboratory; at the beginning of the program, the results had a mean bias of -2.5 cell/µl (95% CI -22.2; 17.1) and at the end of the program a mean bias of -4.6 cell/µl (95% CI -24.6; 15.4); this was -7.1 cell/µl (95% CI -37.2; 23.0) for laboratories participating in 3 or 4 rounds and -1.5 cell/µl (95% CI -27.1; 24.2) for those participating in 5-7 rounds. All confidence intervals listed above included zero (Table 2 & Figure 1).

There was minimal decrease in the discrepancy between results of the participating laboratories from the beginning (LOA -196.9 cell/µl to +191.8 cell/µl) to the end of the CD4 EQA program (LOA -194.6 cell/µl to +185.4 cell/µl). These changes were more enhanced with increasing participation in the CD4 EQA program; for participation in 3 or 4 rounds LOA were -218.9 cell/µl to 204.8 cell/µl, and for participation in 5-7 rounds LOA were -161.0 cell/µl to +158.1 cell/µl. These LOA were all within the clinically significant cut off of ±250 cells/µl (Table 2 & Figure 2).

The number of outliers decreased from 10 at the beginning to 8 at the end of the CD4 EQA program. At the beginning of the program, only 1 was outside the clinical significant range of ±350 cells/µl and none of the CD4 test results were outside the clinical significant range of ±500 cells/µl. At the end of the program, 2 outliers were outside the clinical significant range of ±500 cells/µl.
± 350 cells/µl and none were outside the clinical significant range of ± 500 cells/µl; both of these outliers were from laboratories that participated in 3 or 4 rounds. All outliers from laboratories that participated in >4 rounds were within both clinically significant ranges (data not shown).

The RMSE showed a decrease in variability of CD4 test results at the field laboratories from those of the reference laboratory from the beginning of the participating rounds to the end with a lesser variability for laboratories participating in more than 4 rounds. The RMSE was lower for participating laboratories at the end of the program (RMSE =94.7) compared to participating laboratories at the beginning of the program (RMSE =96.9). Similarly, the RMSE was lower for laboratories participating in more than 4 rounds (RMSE =79.2) compared to those participating in 3 or 4 rounds (RMSE =105.3). (Table 2)

### Table 2: Bias and LOA* analysis, RSME and CCC of field CD4 results field labs in Western Kenya at the beginning and the end of the KEMRI CD4 EQA program, 2014-2015

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Bias (95% CI)</th>
<th>Limits of agreement</th>
<th>RSME#</th>
<th>CCC(^\uparrow) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Field lab at beginning of program (n=166)</td>
<td>-2.5 (-22.2;17.1)</td>
<td>-196.9;191.8</td>
<td>96.9</td>
<td>93 (90-95)</td>
</tr>
<tr>
<td>Field lab at end of program (n=154)</td>
<td>-4.6 (-24.6;15.4)</td>
<td>-194.6;185.4</td>
<td>94.7</td>
<td>94 (91-95)</td>
</tr>
<tr>
<td>Field lab at end of program (≤4 cycles) (n=86)</td>
<td>-7.1 (-37.2;23.0)</td>
<td>-218.9;204.8</td>
<td>105.3</td>
<td>93 (89-95)</td>
</tr>
<tr>
<td>Field lab at end of program (&gt;4 cycles) (n=68)</td>
<td>-1.5 (-27.1;24.2)</td>
<td>-161.0;158.1</td>
<td>79.2</td>
<td>95 (92-97)</td>
</tr>
</tbody>
</table>

*Bias mean difference between the two measurements i.e., the field and reference laboratory was computed for each sample submitted to measure the level of accuracy and if the confidence interval included zero then there was no evidence of systematic bias

* Limits of agreement= Bias ± 2SD The limits of agreement (LOA), (Bias ± 2SD), measures the precision. The clinically important range of ±250/µl cells was used to describe our LOA based on prior CD4 cut off for ART initiation of CD4 count of 200 cells/µl plus as well a reasonable margin of ±50 cells.

# Root mean squared error (RMSE) was used to assess variability between results from the field laboratories and reference laboratory; lower values of RMSE implied low variability in measurements

\(^\uparrow\) Concordance correlation coefficient was used to determine the reliability of CD4 test results by field laboratories

Figure 1: Bland-Altman plot for CD4 results of laboratories in Western Kenya, 2014-2014, showed decrease in spread of differences between the field and reference laboratories from the beginning to the end of the participation rounds
For CD4 cut-offs ≤ 350 cells/µl at the end of the participating rounds there was an increase in sensitivity from 73% to 83% (a decrease in upward misclassification bias) and specificity from 94% to 95% (a decrease in downward misclassification bias). At the beginning of the rounds, the upward misclassification probability was 27% which decreased to 17% at the end of the program whereas it was 17% and 19% for laboratories participating 3 or 4 rounds and laboratories participating more than 4 rounds respectively.

At a CD4 cut-off of 500 cells/µl, the upward misclassification probability was 22% at the beginning and decreased to 6% at the end of the program, whereas it was 8% for laboratories participating 3 or 4 rounds and 3% for laboratories participating more than 4 rounds. Majority of the CD4 results from the field laboratories were distributed away from both cut-offs of 350 cells/µl and 500 cells/µl reducing the chances of misclassification (Table 3).

**Table 3: Misclassification analysis of field CD4 results by CD4 category for field labs in Western Kenya at the beginning and the end of the KEMRI CD4 EQA program, 2014-2015**

<table>
<thead>
<tr>
<th>CD4 Cells/µl</th>
<th>Measure</th>
<th>Beginning of program (n=166)</th>
<th>End of program (n=154)</th>
<th>≤4 cycles (n=86)</th>
<th>&gt;4 cycles (n=68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;350 Patients with CD4 within ±50 cells of cut-off of 350 cells/µl n (%)</td>
<td>27 (16)</td>
<td>27 (18)</td>
<td>14 (16)</td>
<td>13 (19)</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>73 (95% CI 63-84)</td>
<td>83 (95% CI 73-92)</td>
<td>83 (95% CI 72-95)</td>
<td>81 (95% CI 64-98)</td>
<td></td>
</tr>
<tr>
<td>Specificity</td>
<td>94 (95% CI 89-99)</td>
<td>95 (95% CI 90-99)</td>
<td>95 (95% CI 89-100)</td>
<td>94 (95% CI 87-100)</td>
<td></td>
</tr>
<tr>
<td>&lt;500 Patients with CD4 within ±50 cells of cut-off of 500 cells/µl n (%)</td>
<td>24 (14)</td>
<td>18 (12)</td>
<td>10 (12)</td>
<td>8 (12)</td>
<td></td>
</tr>
</tbody>
</table>
Sensitivity | 88(95% CI 81-94) | 94(95% CI 90-99) | 92 (95% CI 85-99) | 97 (95% CI 92-100)
Specificity | 85(95% CI 76-95) | 86 (95% CI 78-95) | 79 (95% CI 65-93) | 94 (95% CI 85-100)

Figure 2: Linear regression analysis of reference laboratory values vs field laboratories for CD4 results at the beginning and end of the participation rounds, Western Kenya, 2014-2015

Reliability of field test results

There was an increment in coefficient of determination ($R^2$), from 86% to 88% and the beginning and end of CD4 EQA program respectively. Similar results were observed while comparing field laboratories participating in 3 or 4 rounds cycles ($R^2=86\%$) compared to those participating in >4 rounds ($R^2=90\%$) (Figure 1) (Figure 2).

Similarly, the correlation of field laboratories CD4 results to that of reference laboratories increased from 93% to 94% from the beginning to the end of the participating cycles. At the end of the CD4 EQA program, CD4 test results from laboratories that participated in 5 or more cycles were better correlated (CCC=93%) than those that participated in fewer cycles (CCC=95%) (Table 2).

IV. DISCUSSION

A. Re-statement of Objective Statement and main results

To the best of our knowledge, this is the first lab regional CD4 EQA program in Kenya that evaluates the accuracy of CD test results generated from field laboratories by repeat testing of the same sample at a reference laboratory. We found that the laboratory results generated at the field laboratories correlated well with those generated at the reference laboratory and similar to results reported in the literature, the correlation improved with an increase in the participation rounds in the CD4 EQA program [5, 11].

B. Reproducibility of field test results

The mean bias generated by our results illustrated that most of the results were within clinically acceptable limits of agreement of $\pm 250$ cells/µl and that there was no evidence of systematic error between the CD4 results at the field and the reference laboratories at the end of the participation rounds [12, 15, 21, 22]. With an overall negative mean bias in the CD4 test results, the CD4 results at field laboratories are therefore likely to be lower than those at the reference laboratories. A decrease in the number of outliers from the beginning to the end of the participation cycles implies a decrease in the variability of CD4 test results. Furthermore, there were no outliers within the clinically significant ranges specified at the end of the participation which were likely to influence clinical decision [23, 24].

www.ijsrp.org
C. Clinical impact on ART initiation

A decrease in both upward and downward misclassification probabilities were observed in our evaluation implying that HIV infected patients would appropriately be instituted on therapy [18, 25].

Although only a slight decrease in specificity with CD4 cut-offs of 350 cells/µl, at the last round of participation with facilities that participated in more rounds, a downward misclassification bias would lead to more patients will be categorized as being in need of ART. Additionally, with the changes in guidelines for initiation of ART to initiate ART at a CD4 count of 500 cells/µl, this CD4 cut-off will no longer be in use [17, 26]. With the move towards ART initiation at higher CD4 counts, misclassification results that end in earlier treatment maybe more acceptable [19]. With more than 80% of CD4 values lying outside the range of ± 50 cells/µl away from the cut-offs, the likelihood of misclassification is limited. However, an ‘acceptable’ margin of error around the clinically important range is still undefined [12, 19].

D. Reliability of field test results

Our analysis also revealed that with an increase in participation, there was an increase in similarity between the results as denoted by the concordance correlation coefficient. Similarly, the likelihood of CD4 results from participating laboratories being similar to that generated by the reference laboratory increased with increasing participation as denoted by the coefficient of determination. With increasing participation in CD4 EQA programs, it is evident CD4 results from field laboratories can therefore be relied upon during patient management.

E. Limitations

We were unable to account for variability in CD4 counts that may have occurred due to technical and physiological reasons [9]. Our analysis is based on reference standards which we presume are accurate; this however, may not always be the case [19]. We were unable to illustrate similar results from laboratories that were excluded from the analysis and yet had more networked peripheral laboratories.

F. Conclusions

Guaranteeing access to quality CD4 testing services in order to support the scale up of ART provision is possible through participation in affordable EQA programs which illustrate the quality of CD4 test results at field laboratories [9]. The costs of participating in UK NEQAS in the year 2015 was USD 250 per laboratory per participating round as compared to approximately USD 100 per round to participate in the regional sample split test CD4 PT per laboratory (Laboratory Manager, KEMRI HIV-R laboratory, Personal communication, 12th January 2016). However, it is by participation in EQA programs that we can illustrate the quality of CD4 test results at field laboratories. Regional EQA programs, which can take into account the existing local infrastructure, logistics and skills requirements can aid in trouble-shooting and providing contextualized solutions to challenges with CD4 testing even in the presence of existing international EQA programs. Validation of CD4 test results should be done regularly and continuously and could provide an affordable alternative to participating in well-established international EQA programs which over time decrease the variability of test results between the field and reference laboratories [7, 10].

ACKNOWLEDGEMENTS

We wish to acknowledge the KEMRI Director, CDC DGHA, all participating laboratories from Ministry of Health, Faith-Based, Non-governmental and Private Health facilities for participating in the CD4 EQA program.

REFERENCES


AUTHORS

First Author – Burmen BMBChB MPH PhDs, Biotech MSCs App Epidemiologu and Laboratory Management, Kenya Medical Research Institute, Centers for Global Health Research, Kisumu, Kenya

Second Author – Mutai K MSc App. Stats, Biotech MSCs App Epidemiologu and Laboratory Management, Kenya Medical Research Institute, Centers for Global Health Research, Kisumu, Kenya

Third Author – A Kamire V MA Project Management, Biotech MSCs App Epidemiologu and Laboratory Management, Kenya Medical Research Institute, Centers for Global Health Research, Kisumu, Kenya

Fourth Author – Frodwa N BSc, Biotech MSCs App Epidemiologu and Laboratory Management, Kenya Medical Research Institute, Centers for Global Health Research, Kisumu, Kenya